

A bid for scientific greatness

The historic dependence of Japan on other people's ideas has served it well, but cannot continue. The government's determination to boost fundamental research seems likely to encourage a cultural shift, despite the obstacles.

JAPAN has done something remarkable. At a time when science in many other countries is under increasing pressure to deliver applications or to meet strategic socio-economic objectives, the Japanese government has committed itself to substantial across-the-board increases in support for basic research in its universities and government research institutes (see page 105). The fact that the Ministry of Finance has accepted the planned 50 per cent increase in government outlays for science and technology to ¥17,000 billion (US\$155 billion) over the next five years is little short of astonishing, given the relatively poor state of the nation's finances. Is it all too good to be true?

Where the money will come from remains to be decided. But the signs are that funding for public works projects, which helped to prop up the former regime of the Liberal Democratic Party for 38 years, will be reduced, and next year's increase in the consumption tax (from 3 per cent to 5 per cent) will also help. Moreover, past experience with long-term government research projects suggests that the government is more likely than most to live up to its commitment. Some caution is due, given that the finance ministry bitterly resisted setting the target. It may still try to find some devious way to boost the figures temporarily — perhaps with one-off supplementary budgets of the kind seen in recent years — rather than provide a sustained growth of support.

But the Science and Technology Agency clearly expects a long-term and substantial improvement in its budget with its extraordinarily ambitious plan to spend ¥2,000 billion on brain research over the next 20 years, or ten times the present level of funding. No doubt other science-related ministries and agencies will deliver equally ambitious plans over the next few weeks as the time for next year's budget proposals at the end of August approaches.

A cynical Western view of all this would be that Japan is at last going to start paying its fair dues and make a contribution to basic research proportional to its economic might. That view is justified: a cool look at the figures shows that Japan is aiming only to bring its spending up to a level comparable to that of advanced Western nations. But the growth required suggests that Japan's publicly funded researchers can look forward to a more fulfilling climate than that facing many of their Western colleagues.

Japan's policy-makers want to boost all areas of science, not only to strengthen the country's future economic base but also to energize its fundamental innovative capacity. They see that although Japan achieved economic success with its cars, electronic consumer goods and computer chips in the 1970s and 1980s, the country has now been left behind in many newly emerging areas of technology, such as biotechnology, gene therapy and software for computer networks, that came out of innovative basic research laboratories of the West. They lament Japan's lack of Nobel prizes and are concerned that the present education system in both schools and universities stifles creative thinking, and they worry that young people are turning their backs on careers in science and engineering. Their aim is to invigorate the university and government research system as a whole and not just to target strategic areas.

It will take a lot more than mere money to achieve this. The five-year plan, for example, also calls for dramatic reforms in employment practices to introduce limited-term contracts for gov-

ernment researchers instead of always guaranteeing lifetime employment. Accompanying this will be more assessment of individual scientists and expansion of competitive research grants, expanded employment of foreign scientists, increased numbers of technicians, and a doubling in the number of postdoctoral fellowships. There is no doubt that extensive changes lie ahead, although pockets of dead wood and inbreeding will inevitably persist in research departments for quite some time.

One important element is still missing. Japan does not have a government body that can take an independent view of the various science-related ministries and agencies. But moves are afoot to establish such an organization modelled on the former US Office of Technology Assessment (see *Nature* 378, 657; 1995). The commitment of Japanese ministries to their plans can be excessive, leading to wasted expenditure. The new independent body is needed to ensure that extra funds for science are sensibly invested.

Crisis not yet over

Boris Yeltsin's election victory gives Russians a chance, eventually, to believe what they are told.

SHORTLY after the first round of Russia's presidential elections in June, Victor Chernomyrdin, the prime minister, spoke at an awards ceremony at the All-Russian Institute of Experimental Physics in Sarov. His message was reassuring. Following a decree signed two weeks earlier by President Boris Yeltsin, spending on science would be increased to 3 per cent of the federal budget, and all money owed to centres such as that at Sarov would be paid. "The crisis has passed," Chernomyrdin said.

Few Russians these days are prepared to believe much of what they hear from politicians. Nor will many believe that Chernomyrdin's statement was unrelated to his desire to boost Yeltsin's second-round prospects. Nevertheless, there are reasons for Russian scientists to find an element of reassurance in the prime minister's explicit promise of increased support — just as they should, in principle, be relieved by Yeltsin's subsequent victory last week.

Many will have been tempted to vote for Yeltsin's main rival, Gennadi Zyuganov, reflecting their dissatisfaction with the situation facing the scientific community. Over the past four years, many government-financed laboratories have been closed. To the highly qualified researchers who have lost their jobs, a return to the security of the past is understandably attractive.

But there are signs that things are improving. Earlier this year, for example, Boris Saltykov, the minister for science and technology policy, was able to announce that, as a result of salary increases introduced last autumn, the gap between such salaries and the official subsistence level has, at least, stopped increasing.

The difficulties facing Russian science, like the rest of the economy, remain enormous. But at least Yeltsin's victory has left open the doors for reform, including the continued need to increase the flexibility and transparency with which Russian science is run. The West can only hope that his success will have given Yeltsin and his allies the confidence to grasp the nettles that they still face. □



US company comes under fire over patent on umbilical cord cells

Paris. Research groups from France, the Netherlands and the United Kingdom are planning to challenge a European patent on the use of stored stem cells from umbilical cord blood that was granted in May to the US company Biocyte Corporation. Such cells are widely thought to hold considerable therapeutic promise for bone marrow transplantation and gene therapy.

Biocyte's patent covers "hematopoietic stem and progenitor cells of neonatal or fetal blood, that are cryopreserved, and the therapeutic uses of such stem and progenitor cells upon thawing". The patent would give Biocyte rights over making, keeping, selling, disposing, importing and exporting stem cells from cord blood, as well as their therapeutic uses.

The research groups opposing the patent are all members of Eurocord, the European Cord Blood Bank which involves 14 European teams. Acting on behalf of the European Group for Blood and Marrow Transplantation, they say their main opposition to the patent is on moral grounds.

In the past, researchers have held back from patenting the various sources of stem cells, such as bone marrow, fetal liver, thymus and spleen, that can be used for transplantation. "We strongly feel that cord blood should not be patented," says Frederik Falkenburg, of the Department of Haematology at Leiden University Hospital in the Netherlands. "We believe it is not ethical to patent a human tissue."

Another critic is Eliane Gluckmann, a researcher at the Hôpital St Louis in Paris, who says that the patent contravenes a resolution approved by the International Society of Transplantation that parts of the human body should not be commercialized, and that organ donations should be free (see page 108). Others are concerned that the potential threat of expensive litigation may discourage companies and research groups from exploring new uses for cord blood cells.

Gluckmann says that Biocyte's patent could, in principle, allow it to prevent cord blood banks from operating, or at least require them to pay licences. Jean Dausset, Nobel laureate for physiology or medicine in 1980, says he is "scandalized" by industry's "seizure" of bone marrow transplants.

Gluckmann is leading the European campaign to challenge the patent. In 1988, working in collaboration with the patent holders, her group was the first to successfully use stem cells from umbilical cord in bone marrow transplants, transfusing a child with Fanconi's anaemia (an often-fatal

genetic disorder that affects red blood cells) with cord blood from its unaffected sibling.

Rodman Rockefeller, a director of Biocyte, confirms that the company is keen to secure licenses world-wide for its patented technology. "We feel it is an important discovery that will become more so over time," he says, adding that any opposition to the European patent will be aggressively challenged. But he says Biocyte is "willing to license its rights under its



Gluckmann: leading challenge to patent.

European patent", and would not prevent cord blood banks from operating.

Such banks have already been set up in many countries. Although only about 200 cord-blood transplants have so far been carried out — compared to the 21,000 bone marrow transplants carried out in 1994 alone — physicians believe that the technique holds considerable promise.

Cord blood is a valuable source of stem cells, which can produce red and white blood cells, as well as platelets. Transplants of these stem cells are widely considered to be more effective than the conventional practice of taking stem cells from bone marrow donors, which is expensive, and painful for donors.

ViaCord, the leading US commercial cord blood bank, estimates that 9,000 patients die each year waiting on a bone marrow transplant. In contrast, umbilical cord blood transplants can be stored frozen, and are available immediately off the shelf.

Cord-blood stem cells also provoke less of an immune response, considerably reducing the risk of graft versus host disease. The technique could be used to treat cancers and other diseases, such as leukaemia, advanced solid stage tumours and genetic disorders, through bone marrow transplantation and gene therapy.

Moreover, genes seem to be easier to introduce into cord blood stem cells than into those from other sources, while introduced genes would in theory be expressed in all their blood cell progeny. The US National Institutes of Health has invested \$37 million on research into cord-blood transplantation, and plans to establish four cord-blood banks.

The European groups say that their

main complaint is that the patent is unethical. But they also believe that it does not meet the legal criteria of novelty and 'non-obviousness'.

Falkenburg says he is confident that the critics of the patent will be able to show that it does not describe a novel invention. The use of cord blood, he asserts, "is not an invention by one group, but a logical development [of research carried out by several groups]; the idea was not new".

In the United States, Biocyte's patent was challenged in 1993 by Cryo-Cell International, a company based in New York. This led the US patent office to open a re-examination in the same year. Biocyte has appealed this decision and no final judgement has yet been made. Cynthia Fisher, president and chief executive officer of ViaCord, says that Biocyte's technology consists merely of "known cryopreservation techniques that have been utilized for bone marrow for decades".

Over the past few weeks, the European researchers have been trawling through the medical literature for evidence to support their case. They claim that various papers dating back to 1977 describe similar techniques to the patent.

But the paper that may clinch the case, says Falkenburg, is a Spanish PhD thesis, published in 1985, which explicitly states, under the heading "Potential clinical uses of cord blood", that cord blood contains sufficient hematopoietic cells to carry out a transplant. "We think this will show that the patent is not inventive", says Falkenburg.

Adriane Antler, a lawyer with Pennie & Edmonds, the firm representing Biocyte, however, dismisses such claims. "People are now jumping on the bandwagon," she says. But the patented process, she adds, "was not evident at all at the time of the filing of the patent".

Declan Butler

'Relief' in Russia

Moscow. "Most Russian scientists will have given a sigh of relief that our country will continue to develop in a direction chosen by its people," says Boris Saltykov, Russia's minister of science and technology, referring to last week's re-election of president Boris Yeltsin.

Speaking in Moscow on Monday, Saltykov said the main result of the presidential election was that "we are now able to plan our activities for years ahead, rather than speculate on what could happen in a month or two". □

Roche fails to get European patent on *Taq* polymerase

London. The European Patent Office (EPO) has rejected a claim from pharmaceutical company Hoffman-La Roche for a patent on the naturally occurring version of the thermostable enzyme *Taq* polymerase, used in laboratories throughout the world in the polymerase chain reaction (PCR) process to amplify small amounts of DNA.

In a letter to the company, the patent office suggests it can find no significant difference between the *Taq* described in the patent application, which was first submitted in 1987, and that resulting from experiments carried out at the University of Cincinnati in 1976 (see *Nature* 373, 377; 1995).

Some observers believe that the decision, if it survives an appeal, could eventually lead to a significant fall in the cost of *Taq* throughout Europe as, although the enzyme is widely believed to cost less than 20 US cents a unit to produce, suppliers which have signed a licensing deal with Roche are said to have to pay a further 16 cents to the Swiss-based company.

It could also have an impact in the United States. Although *Taq* has been patented there by Roche, this patent is already being challenged by the laboratory supply company Promega, in a case which is due to go to trial in San Francisco in the near future.

Carolyn Sutter, a spokeswoman for Roche, says that the company does not see any necessary connection with the US case, which follows a charge brought by the company against Promega for unlicensed sales of *Taq*. "It is a completely different legal situation" in the United States to that in Europe, she says.

Sutter says that Roche is planning to challenge the arguments put forward by the EPO examiner in rejecting the European patent application, and is discussing the next step with its lawyers. "We believe in the patentability of this invention," she says.

But Bill Linton, chief executive officer of Promega, disagrees. "The EPO has recognized the strength of the 1976 prior art in denying Roche's nine-year attempt to obtain the *nTaq* patent," he says. "The US District Court in San Francisco is looking at the same evidence, and we believe it will reach the same conclusion — *nTaq* should not have been patented."

Roche obtained the rights to the potential *Taq* patent as part of a \$300-million deal in 1991 with US company Cetus for the rights to PCR, a technique invented by the Nobel laureate Kary Mullis that has revolutionized the analysis of DNA samples.

Although the validity of the PCR patent has not been challenged, the *Taq* patent has long been controversial, both because of licensing fees reported as being demanded

by Roche, and also because *Taq* is a naturally occurring enzyme extracted from the bacterium *Thermus aquaticus* found living in hot springs (the company already has a patent on recombinant *Taq*, which is not under dispute).

Roche has actively pursued companies discovered to be selling *Taq* without a licence, and has persuaded large users of the enzyme, such as Britain's Medical Research Council (MRC) and its Biotechnology and Biological Sciences Research Council, only to deal with licensed suppliers, at least until the patent issue is settled.

Two years ago, in what now turns out to have been a premature move, Roche announced that it had been told by the patent office that the patent would be granted. But its confidence was misplaced. The examiner subsequently notified the company that, in view of objections raised



(one of the most prominent challengers being Promega), it required further evidence that the *Taq* described in the patent application had not already been discovered (see *Nature* 374, 108; 1995).

The company's response was that, first, the Cincinnati scientists had described an enzyme with molecular weight significantly different from its own and, second, that Roche's enzyme was much more effective in faithfully duplicating DNA sequences. The patent examiner has not been convinced, and has informed Roche that her original objections to the patent "still apply".

The next legal step open to Roche is to appeal against this decision to a three-member appeals board set up by the EPO. If this overturns the examiner's ruling, then the patent will be granted. If the appeal fails — and even if companies currently producing *Taq* under licence have difficulty in changing the terms of their agreements with Roche — others will be free to sell products based on naturally occurring *Taq* without paying royalties.

David Dickson

US panel established to coordinate work on endocrine disrupters

Washington. The Clinton administration has established a panel to improve the coordination of research into the effects of endocrine disrupters, the man-made chemicals whose presence in the environment may be disrupting human reproduction and development by interfering with hormone levels.

The high-level panel — a working group of the environmental subcommittee of President Bill Clinton's National Science and Technology Council (NSTC) — will bring together scientists from a dozen federal agencies to establish the scope of existing work on endocrine disrupters, and to fill gaps in the programme.

The panel starts its work amid considerable controversy in the United States about the publication of a book, *Our Stolen Future*, by Theo Colburn and Dianne Dumanoski, which argues that the chemicals constitute an urgent public health threat. The book, which has a foreword by Vice-President Al Gore, has been attacked by industry groups and scientists who say that the threat is, at best, unproven.

But Bob Kavlock, a panel member and a leading toxicologist at the Environmental Protection Agency (EPA), says the research is justified, whether the threat is proven or not, because its potential impact is so serious. "If it affects reproduction, it has fairly big implications," says Kavlock, citing the current scientific dispute about sperm counts in American men as an example. "If human sperm counts really have declined by half over the last 40 years, we need to know about it now — we can't wait another 40 years to find out," he says.

Kavlock, who is director of the reproductive toxicology division at the EPA's National Health and Environmental Effects Research Laboratory at Research Triangle Park, North Carolina, says that the panel is starting work by developing a summary of research needs in key areas. These include methods to detect the presence of the chemicals and their effects, and models to extrapolate available data to show the likely impact on public health.

The panel will also prepare an inventory of existing programmes in each of the agencies, as well as surveying work under way outside the United States, before assessing how well the work fits research needs.

The EPA currently spends almost \$10 million a year on endocrine disrupter research, which it has been studying (under various names) for 20 years. The size of the total federal effort is unknown, but could well increase as a result of Gore's interest and the activities of the NSTC panel.

Colin Macilwain

'Beware of hype,' AIDS conference told

Vancouver. Leading AIDS scientists warned this week against incautious optimism and excessive hype over early results of new, clinical trials of combination drug treatments for AIDS. Their warning came as the eleventh International Conference on AIDS opened in Canada with a barrage of good news about the war against the disease.

The latest results of several small clinical trials of different combination therapies were due to be released at Vancouver today (11 July), indicating that the therapies are able successfully to repress the viral load of subjects for at least a year.

The trial results will confirm that the expensive and complex drug regimens are able to drive viral load back down below the level detected by blood tests. But there is no evidence that the virus can be eradicated from the body, where it almost certainly remains in the lymph nodes and the central nervous system.

The optimistic mood spurred by these trials among the 15,000 scientists, doctors, activists and journalists who are attending the conference was compounded by Robert Gallo, the veteran AIDS researcher who now directs the Institute of Human Virology in Baltimore, Maryland. Gallo released new evidence of a correlation between levels of chemokines and a natural ability to resist HIV. He predicted that treatment with chemokines or other 'biological' inhibitors would ultimately provide a way of safely controlling HIV.

The United Nations programme on AIDS (UNAIDS) also reported that two of the developing countries hit earliest and hardest by the pandemic — Thailand and Uganda — have successfully reduced the rate of HIV infection, suggesting that health and education programmes are working, even in the most adverse circumstances.

But the domination of the conference by soaring expectations for the combination therapies — which normally put patients on three drugs, including protease inhibitors and the newer reverse transcriptase inhibitors — led officials repeatedly to stress the limitations of such therapies.

"We are at a crossroads in the history of the pandemic," said Martin Schechter, an AIDS researcher at the University of British Columbia and conference co-chair. "We are beginning to see glimmers of hope — but we must stress that they are only glimmers." Peter Piot, director of UNAIDS, warned that the combination therapies may raise unrealizable expectations: "Let's not allow the pendulum to swing from pessimism to hype and excessive optimism."

William Paul, the director of the Office of AIDS Research at the US National Institutes of Health, who directs AIDS research programmes worth \$1.4 billion, warned of a "dreadful duality" if the industrialized

nations pursue costly and sophisticated drug therapies at the expense of approaches that might benefit the poor countries, where 90 per cent of new HIV infections now arise. Paul and others called for renewed efforts to find a protective vaccine to prevent HIV infection (see below).

Results from various small combination therapy trials due to be released at Vancouver this week confirm that the drugs are as successful in suppressing viral load after a year as they had been after six months. In most cases, patients who began therapy shortly after being infected by HIV develop no physical symptoms, and have their viral load reduced to levels undetectable in the blood stream. The trials also uncovered few significant side-effects.

But the therapies cost \$20,000 or so, and as Piot pointed out, even if that fell to \$5,000 it would still cost \$100 billion a year to treat the 20 million infected individuals on the planet. Critics maintain that any deviation from the strict regimens used in the current small, closely monitored trials may foster resistance to the therapy. Concern also surrounds the unknown effects of the toxicity introduced to the body through indefinite treatment with these combinations.

Gallo, meanwhile, reported that unpublished data from a study of haemophiliacs by an Italian team led by Daniel Zagury, and

from a British study of homosexual men, both show strong correlations between the level of C-chemokines in the body and its resistance to HIV in the face of direct exposure. He told a packed hall, with many hundreds locked outside, that this and other biological therapies could be cheaper and far less toxic than the chemical inhibitors now being used in combination therapy.

Attempts to control the expansion of AIDS in the developing world by changing public behaviour were bolstered by news from Uganda, where HIV prevalence among pregnant women had fallen from 21 per cent in 1990–93 to 15 per cent in 1994–95. In Thailand, prevalence among male army conscripts has fallen from a peak of 3.7 per cent to 2.4 per cent last year, and is still falling, according to Damrong Boonyoen, director of the Thai Department of Communicable Disease Control.

Drug companies kept a high profile throughout the conference, shipping in doctors from all over the world and running major symposia next to the main programme. Apart from a few noisy protests, most AIDS activists at the conference appeared to welcome the renewed interest of the pharmaceutical companies — in sharp contrast to the mutual fear and loathing expressed at some previous AIDS meetings.

Colin Macilwain

Double push for new vaccine candidates

THE growing realization that AIDS vaccine development work has virtually ground to a halt (see *Nature* **378**, 323; 1995) led to two major initiatives being announced at the International Conference on AIDS this week, each aimed at kick-starting an effective programme.

William Paul, director of the Office of AIDS Research (OAR) at the US National Institutes of Health (NIH), promised to "increase resources for research on vaccines" and to launch what he called a "restructured, redirected vaccine research programme" at NIH.

Meanwhile, a newly established International AIDS Vaccine Initiative (IAVA), backed by start-up funds from the Rockefeller Foundation and other charities, pledged to raise \$50 million over the next three years to help to prepare vaccines for use in poor countries.

Paul said that he would fully implement the recommendations of the vaccine component of a recent comprehensive review of the AIDS research programme, which calls for a larger, fully integrated vaccine research programme led by "a distinguished non-governmental scientist".

But Seth Berkley of Rockefeller Foundation, who chairs IAVA, said that the current low level of NIH support for vaccine development is "scandalous". Although NIH spends 8 per cent of its AIDS budget — around \$100 million a year — on vaccine work, only about \$5 million is being spent worldwide on preparing new candidate vaccines, Berkley says. Industry has stopped supporting such work because of diminishing prospects of large-scale vaccine trials in the United States, he adds, and the NIH has not filled the gap.

IAVA has \$5 million for its first year, but hopes to raise between \$8 and \$12 million next year, and \$20 to \$30 million the year after that, says Berkley. Peggy Johnston, IAVA's scientific director, says that some of money might come from the World Bank.

Johnston, who was deputy director of the Division of AIDS at the National Institute of Allergy and Infectious Diseases (NIAID) before joining IAVA this year, says that NIH has been relying on industry to come up with potential vaccines, and that "a critical gap" has developed, which IAVA can help to fill.

C. M.

Cluster relaunch looks to minisatellites

Munich. The single back-up satellite of the ill-fated Cluster mission, whose four satellites were on board the Ariane-5 launcher that exploded during its maiden flight last month, is to be prepared for a relaunch. But a decision about when this should take place — and whether the back-up unit should be accompanied by other satellites to allow the major scientific goals of the mission to be achieved — has been deferred to November.

Meeting in London last week, the Science Programme Committee (SPC) of the European Space Agency (ESA) said that a quick launch of the spare satellite, which has been named Phoenix in reference to the ashes from which it will rise, would at least save ESA's commitment to the International Solar Terrestrial Programme, a series of missions flying in different orbits (see *Nature* **381**, 721; 1996). Cluster was intended to take measurements in one of these orbits.

Because of its relatively low cost, estimated at ECU30 million (US\$37 million), this decision will have no impact on ESA's long-term science programme. The SPC, which is made up of delegates from ESA's 14 member states, has decided it needs more time to tackle the thornier — and much more expensive — problem of what steps could be taken to save the other scientific aims of the mission.

Four satellites flying in fixed formation — which was Cluster's unique scientific feature — would be required for the three-dimensional analysis of the Earth's electric and magnetic fields as they interact with solar wind particles (see *Nature* **381**, 541; 1996).

The SPC has rejected the request of French scientists to replace exactly the large and identical satellites which comprised Cluster, and has instead asked ESA's executive and its member states to put together cheaper proposals for three additional minisatellites.

These proposals will be presented to the next SPC meeting in November. But the question of how they would be financed — if they are approved — remains.

ESA has not ruled out the possibility of using its own science budget to pay for them. If it were to do so, however, it would certainly have an undesirable knock-on effect on its long-term science programme, Horizons 2000. The SPC has now delayed approval of the plans for the next stage of this programme until November, in order to accommodate any possible delays resulting from a Cluster rescue plan.

But ESA's clearly preferred plan would be for participating member states, which normally pay only for scientific instruments, in this case to pay also for the flight units. Despite tight national budgets in European countries, this is "not a pipe dream", says John Credland, Cluster's project manager.

Several member states that already have

minisatellite programmes, including Britain, Germany, Sweden and Denmark, have expressed interest, he says. Mike Blackwell, a spokesman for the British National Space Centre, says that it is far too early to say if interest will be turned into cash. "But there is at least a positive mood in the UK" for such an investment, he adds.

The launch date for Phoenix could depend on whether such a multisatellite mission is approved. The predominant view in the SPC is that it should be flown as soon as possible, to avoid losing any further time in gathering data. Multipoint recording would then begin as soon as Phoenix is joined by its companion flight units.

But some members of the committee fear that, in such a case, the instruments on Phoenix could deteriorate before the new minisatellites reach orbit. They argue that, if ESA agrees to a four-satellite mission, the satellites should all be launched together.

Credland replies that the instruments are more likely to deteriorate on the ground than in space, where there is no polluting atmosphere. A bigger problem, he says, will be coordinating the preparatory work on four satellites that would almost certainly be produced in different countries.

As well as the launch date, it will also be decided in November whether to fly Phoenix free on the next Ariane-5 qualification flight, whose timing is likely to coincide with the completion of Phoenix next spring. This is the option favoured by ESA's space

science directorate, but this time the directorate will argue for closer monitoring of the review process that assesses the launcher's readiness for flight. Last time it was brought into the process only at a late stage.

A report on the investigation of the cause of Ariane-5's failure will be published next week. It is expected to put the blame on a crucial piece of software in the control of flight direction, used successfully on the

much smaller Ariane 4. ESA officials are known to consider that this type of fault should have been picked up at an early stage.

The SPC also learnt last week that plans to restructure ESA, requested by Europe's space ministers last November as a means of increasing efficiency, have been modified. This followed objections from David Southwood,

professor of physics at Imperial College, London, and chairman of the committee, who had argued that the proposed separation of planning and implementation in ESA's science programme — the only mandatory activities of the agency — would jeopardize the reliability of long-term planning (see *Nature* **381**, 353; 1996).

Jean-Marie Luton, ESA's director general, is now suggesting an interim plan, under which the whole of the space science directorate would come under the larger umbrella of a new planning directorate, to include all of ESA's activities — including Earth observation, telecommunications and utilization of the international space station.

Alison Abbott



Fallen to Earth: remains of Cluster in the swamps of French Guyana. Some of the debris was in much better condition than expected, and some electronic parts from the solid-state recorders may be retrievable for possible reuse.

India appoints 'part-time' science minister

New Delhi. The appointment by India's newly elected United Front government of a part-time minister to look after the country's huge science and technology enterprise has disillusioned the scientific community. Indian scientists had been expecting the new government to give greater priority to science than did the previous Congress government, and now feel let down.

The new minister is 57-year-old Yoginder K. Alagh, an economist trained in the United States who was vice-chancellor of the Jawaharlal Nehru University in New Delhi before he joined the council of ministers last week. As a minister of state without cabinet rank, Alagh is primarily responsible for planning and programme implementation. He has been given the science ministry as an "additional charge".

"Non-appointment of a full-time minister is a clear sign that science is not high in the agenda of the new government," says a senior official in the science ministry. He says it is sad because science did not get the attention it deserved even when it had a full-time science minister in the Congress government.

In the past five years, universities and publicly funded laboratories have been run on a shoestring budget. In real terms, funding, as a percentage of gross domestic product, dropped from 0.93 per cent in 1989 to 0.83 per cent last year. Of the 70 billion rupees (US\$2 billion) spent on research and development in 1995, the central government's contribution was only 50 per cent — the rest came from the states and industry.

K. S. Jayaraman

Onshore dismantling 'the best option' for most oil platforms

London & Munich. Disused oil installations made from large quantities of steel should be dismantled and recycled on land, and not dumped in the sea, according to a report commissioned by the German government.

The report, *Regulations for Decommissioning Offshore Platforms in the North Sea and North Atlantic*, a summary of which is due to be published in Berlin this week, says that onshore dismantling of steel installations is healthier for the marine environment and cheaper, as 95 per cent of most structures can be recycled for scrap metal.

But the report recommends that concrete structures should be sunk, because they cannot be recycled. Their prohibitive weight — around 1 million tonnes — rules out lifting by cranes or towing with barges. Concrete structures represent 5 per cent of the 400 oil installations in the North Sea; the rest are made from steel.

The findings have been welcomed by the environmentalist group Greenpeace, which last year succeeded in preventing the Shell oil company from dumping Brent Spar, a 14,500-tonne disused steel oil-storage buoy, in the North Sea (see *Nature* 376, 378; 1995). But they have been rejected as "irrelevant" by Shell, which last week announced a 21-strong list of bids from consortia to decommission the buoy, currently moored off the Norwegian coast.

Shell reversed its original plan following a public outcry, particularly in Germany. Critics opposed the dumping on the — controversial — grounds that the structure could harm the deep-sea environment.

The company now says that it can no longer rule out deep-water disposal, as a new structural survey of Brent Spar suggests that the buoy could break up while being towed to land.

But Markus Grosse Ophoff, a scientific adviser to Germany's Federal Environment Agency, which compiled the report, says that oil installations need not be towed in once piece. "Oil installations are built in sections," he says. The engineering capability exists, he adds, to separate the sections offshore before towing them to land.

However, Dave Stuart, a spokesman for Shell, says any attempt to bring Brent Spar on to land will require raising the 137-metre-high structure further out of the water; most of Brent Spar is submerged. The structural survey shows that the walls of Brent Spar's storage tanks have been weakened and might collapse during any attempt to pump water out of the tanks — to raise Brent Spar. Stuart reiterates, however, that Shell is not ruling out any option.

Industrial world is unlikely to meet promised carbon cuts

London. Commitments by most industrialized countries to reduce emissions from fossil fuels to 1990 levels by the end of the decade are unlikely to be met, according to a report published last week by the World Energy Council (WEC), an international body supported by the energy industry.

Global carbon dioxide (CO₂) emissions from burning fossil fuels rose by 12 per cent in the first half of the current decade, says the report. The largest rise has come from developing countries, whose increase in emissions from fossil-fuel burning has been more than three times the increase for industrialized countries.

But the survey also shows that industrialized countries — with the exception of France, Germany and the United Kingdom — are unlikely to meet the goal of returning CO₂ emissions to 1990 levels before the start of the next century, as agreed under the United Nations (UN) climate convention signed in Rio de Janeiro three years ago. In contrast, CO₂ emissions in eastern Europe and the former Soviet Union have already dropped to below their 1990 levels.

The survey was published to coincide with this week's UN meeting — the second annual conference of the parties to the climate convention — one of whose goals is to negotiate new targets to reduce CO₂ emissions beyond 2000.

The WEC survey shows that between 1990 and 1995, most member countries of the Organization for Economic Cooperation and Development (OECD) registered a 4 per cent rise in CO₂ emissions. But countries in the Asia-Pacific region — excluding Australia, New Zealand and Japan — showed a 30 per cent rise. Emissions in countries in the Middle East increased by 35 per cent, in Africa by 12.5 per cent and in Latin America by 8 per cent.

Economic decline contributed to a reduction in CO₂ emissions in countries of the former Soviet Union and central and

eastern Europe, according to the survey. In 1995, emissions from the Commonwealth of Independent States (CIS) of the former Soviet Union were 70 per cent of their 1990 levels, and 75 per cent for central and eastern Europe. But the combined CO₂ emissions for both regions in 1990–95 increased by 3 per cent.

Countries in the developing world are under no obligation to reduce their CO₂ emissions, as they are not party to the climate convention. On the basis of the WEC's statistics, future draft agreements on reducing emissions are unlikely to exclude them. However, many developing countries consider cutting back on CO₂ emissions an unwelcome brake on their drive for industrialization.

"Reducing CO₂ emissions because of climate change has placed developing countries in a really serious bind," says Abdul Bar al Gain, director of Saudi Arabia's environment agency, and head of the country's delegation to the UN climate talks.

"It's like telling a poor trader who saved for 20 years to buy a van to carry on making deliveries using his donkey cart. It's not fair. We need to find some other solution for those developing countries that will become increasingly dependent on fossil fuel burning."

Anwar Ibrahim, deputy prime minister of Malaysia, points out that industrialized countries continue to be responsible for more than half of the world's CO₂ emissions. But he says that both sides need to be flexible in negotiations on future emissions reductions.

"It is very important not to condemn the West," he says. "Otherwise this will give licence to developing countries to rape their forests and pollute their environment." Carbon dioxide emissions from burning oil in Malaysia increased by 66 per cent between 1990 and 1995.

Governments are not required under the convention to agree on steps towards further emissions reductions beyond 2000 until the third annual climate-convention conference, due to be held next year. Michael Jefferson, deputy secretary general of the WEC, says that, on the basis of this latest survey, setting firm targets for further reductions would not be realistic at this stage.

But Helen Wallace, a senior scientist with Greenpeace, says the survey will provide a stronger focus for calls to reduce CO₂ emissions to 20 per cent of their 1990 levels by 2005. Environmental groups — including Greenpeace — belonging to the Climate Action Network regard the focus on developing countries as "a dangerous and time wasting" digression.

Ehsan Masood

CO₂ emissions from fossil fuel combustion (Gigatonnes of carbon)

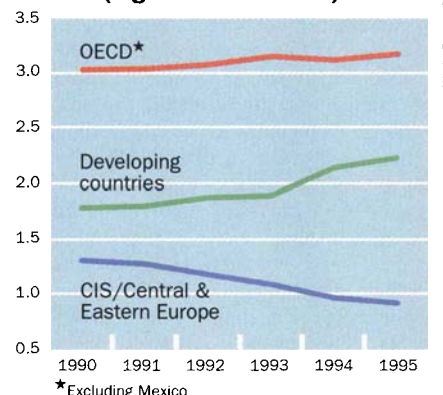


IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

VentureStar: government funding is planned to cover initial development costs only.

NASA aims high for reusable spacecraft

Washington. The National Aeronautics and Space Administration (NASA) last week chose the Lockheed-Martin Corporation to build and test a prototype reusable launch vehicle (RLV) that it hopes will dramatically reduce the cost of reaching orbit, beginning around 2005.

The company's VentureStar design beat two other contenders for NASA's X-33 test programme, which will build and fly a half-scale model of the RLV. The vehicle will take off vertically and land horizontally like an aeroplane, and will have no pilot or passengers on board.

NASA has budgeted \$941 million for building the X-33 and conducting at least 15 test flights in 1999. Lockheed-Martin will contribute a further \$220 million. After that, however, the government intends to step out and let industry finance, build and operate a working vehicle.

Daniel Goldin, the administrator of NASA, said that the RLV programme aims to cut the price of delivering cargo to orbit to only \$1,000 a pound, or more than ten times below current costs. Most outside analysts believe this is too optimistic, but they say the RLV has a better chance of being economical than the current space shuttle, which never came close to achieving its original cost projections.

One of the RLV's missions would be to re-supply the international space station. Beyond that, the market for such a vehicle is uncertain. The RLV could have a difficult time competing against single-use rockets for the satellite launch market, and NASA would probably have to make a commitment to buy a large number of flights up front before any private company would be prepared to agree to build an operational version of the launch vehicle.

But the commercial viability of the new vehicle could suffer if it ends up carrying humans. A NASA advisory group last week warned the agency to slow its plans to cut the current shuttle budget, or risk compromising crew safety.

Tony Reichhardt

German university withdraws doctorate after fraud charge

Munich. In what could be the first case of its kind in Germany, a chemistry researcher at the University of Bonn has had his doctorate withdrawn following charges of scientific fraud. But the researcher, Guido Zadel, maintains his innocence, and is likely to challenge the decision in court.

Zadel, a former doctoral student at the university's Institute of Organic Chemistry, was awarded his PhD at the end of 1993. But it was rescinded by the faculty of mathematical and natural sciences earlier this year, two years after he had been caught trying to manipulate the results of a laboratory experiment based on his thesis, which had described a technique for distinguishing 'left-handed' from 'right-handed' molecules.

Through his lawyer, Zadel acknowledges that he committed a single "stupid" act. But he says it was prompted by pressure to obtain positive results, as no group has been able successfully to reproduce his thesis results, which he still defends. The faculty, however, took a less sanguine view. In a 53-page report issued in April, it accused him of scientific fraud.

Zadel has claimed to have made a highly significant discovery in the area of 'chiral synthesis', a field of considerable interest to the pharmaceutical industry. So-called chiral molecules are known as enantiomers and exist in two chemically identical forms: left-handed and right-handed molecules.

Pharmaceutical products need to be tested separately for the effects of left- and right-handed molecules, following the tragedy in the early 1960s when women gave birth to children with malformed limbs after being given the sedative thalidomide, which had been contaminated with the wrong enantiomer.

Current testing techniques, however, are expensive and complicated. So when Zadel claimed that the application of a static magnetic field during a chemical reaction would produce an excess of left- or right-handed enantiomers, there was wide interest. He even claims to have discovered that the size of the enantiomer surplus varied according to the size of the magnetic field applied.

On the strength of this unusual yet potentially important result, the German research council, the Deutsche Forschungsgemeinschaft (DFG), extended funding for the project to allow the results to be verified, despite the lack of any theoretical basis for the observation.

The work was published in *Angewandte Chemie*, the German chemistry journal, in April 1994 (see *Angew. Chem.* **106**, 460; 1994). That, however, was when the trouble started for Zadel and his supervisor Eberhard Breitmaier, a professor at the Institute

of Organic Chemistry, because no other research group could reproduce the results.

Zadel was first accused of fraud in May 1994 by two other students of Breitmaier. Claiming now that he was "under pressure", Zadel admitted to mixing some optically active substances with the initial chemicals.

He later revoked this confession, claiming to have been working with a special method that consisted of 'vaccinating' his chemicals with tiny amounts of enantiomers, though not enough to produce the excess later measured. The vaccination might have looked as though he had manipulated the probes, he says.

But Breitmaier, who at first supported Zadel, began to lose faith, and decided to try to catch him out. One of his colleagues set up the experiment again. When the colleague left the room briefly, Zadel replaced the samples with his own enantiomeric mixtures, not knowing that his colleague had hidden half of the test-tubes. Zadel now admits this act was "stupid". The evidence against Zadel seemed incontrovertible and Breitmaier withdrew the paper last July.

But Zadel continues to insist that his results are legitimate and reproducible. Through his lawyer, Johannes Latz, he claims that proof can be found in his laboratory notebooks, which he says were stolen. In addition, he cites successful experiments carried out by a colleague while he was on holiday, as well as experiments performed in front of Breitmaier.

The case has generated considerable interest because allegations of scientific fraud are rare in Germany. Wolfgang Frühwald, head of the DFG, believes this is because German researchers do not have to fight as hard for funding as their colleagues in the United States, where, he says, misconduct is much more frequent.

Frühwald fears that a broader discussion of scientific misconduct might create a climate of mistrust in the community, something he believes should be avoided. Commenting on the Zadel case, Frühwald says the DFG may consider asking the university to return the funds.

But Albin Eser, a director of the Max Planck Institute for Foreign and International Law in Freiburg, says fraud in German science is more widespread than is generally acknowledged. One of Eser's former students, Stefanie Stegemann-Boehl, completed a PhD thesis comparing fraud in biomedical research in Germany and the United States, and concluded on the basis of extensive interviews that many more cases of fraud in Germany may exist than are made public.

Jeanne Rubner

Brain science benefits from budget plan

Tokyo. The potential benefits of a new five-year plan for science and technology put forward last month by Japan's Council for Science and Technology are already becoming apparent. The Science and Technology Agency (STA) has drafted an ambitious plan to invest ¥2,000 billion (US\$18 billion) in brain science over the next 20 years, and more than ¥100 billion in a state-of-the-art nuclear magnetic resonance (NMR) centre for studying the structure of protein molecules.

Both proposals need to be approved by STA's finance section before a request is submitted in late August to the Ministry of Finance, where they will undergo close scrutiny before the national budget is announced in December. But optimists argue that the government's commitment to science funding, which was enshrined in a science and technology law passed last year, will ensure that the plans will be approved with only minor trimming (see *Nature* 378, 227; 1995).

The five-year plan calls for an increase of more than 50 per cent in the government's spending on science and technology, much of it to strengthen basic research in universities and government research laboratories, and was approved by the cabinet (including the Ministry of Finance) late last month.

Some people claim that the plan's emphasis on expanding basic research contrasts with moves in other advanced economies to make research more commercially relevant. But others, such as Keisuke

Toyama of Kyoto Prefectural Medical University, a member of the STA committee that drafted plans for the brain research programme, say it will simply bring Japan's balance between basic and applied research into line with other advanced countries.

The proposed brain science research programme will be organized around a new brain research institute, which will form the hub of a network of designated laboratories chosen from university and government centres that are already engaged in brain-related research. Overseeing the project will be a coordinating committee, which will foster cooperation between the ministries and agencies operating the laboratories.



Ito: masterminded brain science plan.

The committee that drafted the brain science plans is chaired by Masao Ito, president of the Science Council of Japan (JSC) and director general of the Frontier Research Programme of the Institute of Physical and Chemical Research (RIKEN). Ito, who is one of Japan's leading brain researchers, says he is extremely pleased with the plan; "this has been our dream".

Ito says the ambitious budget proposed for this project is a direct result of the government's five-year plan, and that brain research is an area ripe for development. 'It is time to push it forward,' he says. Japan at present spends less than one-tenth of the budget of the US National Institutes of Health (NIH) on brain research, and has only about 3,000 researchers, compared to 25,000 in the United States. The proposed budget, which would amount to ¥100 billion (US\$1 billion) a year, would bring Japan's spending close to that of the NIH.

Ito admits that researchers in other areas of science may regard the planned budget with jealousy. But he insists that life science has in the past been "poorly supported", and is therefore in need of special extra support. "Disciplines such as physics and chemistry have been well supported, but interdisciplinary areas such as brain research have been neglected," he says.

The new brain research institute is likely to be based at RIKEN, as more than half of about 25 frontier research projects there are concerned with brain research. Other laboratories likely to attract funding include the Ministry of International Trade and Industry's Electrotechnical Laboratory and Institute for Bioscience and Human Technology in Tsukuba, north-east of Tokyo, and the Education Ministry's Brain Research Institute at Niigata University, west of the capital. Roughly half the money will go to the central institute and half to

the other laboratories.

The plan, which addresses issues raised by a JSC document on brain research released in May (see *Nature* 381, 266; 1996), sets out three major areas of research: basic research into the structure and function of the brain; medical research aimed at the prevention and treatment of neurological disorders; and research aimed at mimicking the brain's information-processing functions. Forty-seven projects that fit into these categories have already been planned.

Other indications that the five-year plan is producing benefits for science include an STA proposal for an NMR centre for studying protein structures. Initial plans for the centre suggested it would have a budget of ¥50 billion (see *Nature* 381, 105; 1996). But a figure more than twice as large is now being discussed, thanks to the "favourable climate" introduced by the five-year plan, says an STA official in charge of the programme. The agency was due to release detailed research goals for the new structural biology research project on 11 July (see below).

Stephen Barker

Research aims of Japan's new NMR centre for structural biology:

- Elucidation of all basic protein structures.
- Analysis of multi-molecular protein complexes.
- Prediction of the 3D structure of proteins from the amino-acid sequence.
- Determination of the relationships between the higher order structure of proteins and their functions.
- Understand the biological significance of protein structures.
- Manipulate the function of proteins by changing their molecular structure.

Technology to be developed:

- Large-scale X-ray crystallography systems for determining the structure of proteins with molecular weights in excess of 1 million daltons.
- Superconducting magnets with a power of 1 GHz for NMR analysis of large proteins.
- Electron microscope technology for the analysis of membrane bound proteins and other large molecules.
- Neutron scattering crystallography technology.
- Techniques to examine the molecular structure within cells using atomic force and light microscopes.
- Theoretical and computational modelling technology to understand mechanisms of molecular interactions and to determine 3D protein structures purely from genome information.

Finland urges increase in EU research funds

Munich. Finland has become the first member of the European Union (EU) formally to advocate an increase in funds for the European Commission's Fifth Framework Programme for research (FP5), due to be launched in 1998.

In a position paper on FP5, Finland also calls for a greater concentration of funds within fewer programmes, and increasing emphasis on training and mobility programmes for young scientists, as well as on the exploitation of the results of research projects.

All are in line with an emerging consensus among member states about how FP5 might be run more efficiently than the current framework programme.

Finland also suggests that FP5 should be open to non-member industrialized countries to allow the creation of global research consortia. It wants cooperation to be strengthened with the central and eastern European countries that have applied for membership of the EU.

Alison Abbott

Chairman of European assembly will stay on with more staff

Munich. Jan Borgman, former head of the Dutch National Research Council (NWO), last week agreed to remain chairman of the European Science and Technology Assembly (ESTA), the 100-member advisory body to the European Commission established two years ago, following a promise by the commission to increase its administrative support.

Borgmann had announced his intention to resign in March because of the commission's failure to provide adequate support for the heavy burden of work generated by ESTA's increasing number of working groups. The commission has now agreed to provide an extra full-time post in addition to the two full-time and one part-time staff member already devoted to ESTA business. In return, ESTA has agreed to provide more structured plans of its activities. □

Deal for gene-based drugs

Washington. Human Genome Sciences (HGS), the genome-sequencing company based in Rockville, Maryland, has widened its agreement with the pharmaceutical company SmithKline Beecham (SB) to allow several other drug-makers to develop drugs on the basis of genes sequenced and tagged by HGS.

The deal, which was announced last week, will allow Schering-Plough of New Jersey and Synthelabo of France to collaborate with HGS on drug development. The two companies will pay HGS \$90 million over the initial five-year period of the agreements. HGS will also gain the right to select, develop and license half-a-dozen drugs independently from SmithKline Beecham. SB, which has paid HGS \$125 million under the terms of a 1993 deal that bought it exclusive access to HGS gene sequences, will

receive a share of the licence fees and joint rights to manufacture drugs that emerge from HGS's work with its new partners. □

India challenges turmeric patent

New Delhi. India's Council of Scientific and Industrial Research has decided to challenge a US patent on the wound-healing properties of turmeric. The use of turmeric as a traditional medicine is well known in India, and its wound-healing properties were documented by Indian researchers in 1953. The patent was granted to the University of Mississippi Medical Center in Jackson in March. The Indian application for its revocation, on the grounds of 'prior art', is being seen as a test case for many traditional plant-based medicines being patented in the United States. □

Scrapie sheep cull planned

London. Thirty thousand sheep are to be destroyed in Norway in an attempt to eradicate scrapie, the neurodegenerative disease in sheep that is similar to bovine spongiform encephalopathy (BSE) and to Creutzfeldt-Jakob disease, which affects humans.

Scrapie was once widespread in British flocks, and is blamed for the spread of BSE in the United Kingdom, although the bovine disease has never been reported in Norway. A spokesman for Norway's agriculture ministry said that 17 scrapie-infected flocks have already been destroyed. Norwegian farmers have accepted a plan to destroy 600 flocks over two years. □

BSE 'pre-dates rendering changes'

London. The bovine spongiform encephalopathy (BSE) epidemic pre-dates changes to Britain's rendering industry, according to new research published this week. Until now, the UK agriculture ministry has attributed BSE to rendering methods introduced in the 1980s.

To **save an hour**
each time you
purify histidine-
tagged proteins,
put it on the
tip of a syringe



The ministry has said that the new methods allowed the scrapie infective agent in sheep offal to survive and infect cows with BSE via cattle feed. But the results of new experiments funded in part by the European Union, the UK agriculture ministry and the UK's largest rendering company, Prosper De Mulder, shows that the scrapie agent remains active regardless of the choice of rendering method. □

EMBO faces flat budget

Munich. The European Molecular Biology Organization (EMBO), which holds regular conferences and runs a major postdoctoral fellowship programme, had a request for a three per cent increase in funding rejected last week by delegates from its 21 member states, which voted to maintain its current budget, corrected for inflation.

Frank Gannon, EMBO's director, called the decision "very disappointing", as the organization requires more money to maintain the current level of its long-term fellowship programme. There was a "heavy financial cloud" over the meeting, he said, as some member states were apparently unprepared to consider even the "relatively small sums of money that would be involved". □

China's Long March into space

Hong Kong. China restored some confidence in its satellite launch capabilities last week with the successful launch of the Apstar1A communications satellite, a substitute for the Apstar2 satellite which was lost in an explosion during lift-off from China's mountainous launch site in Xinchang, Sichuan Province in January 1995.

The satellite was carried aloft by a Long March 3 rocket, a more tried and tested vehicle than the larger 3B version, which veered off course on its maiden flight in February and is said to have killed a large number of people. At least four foreign launches originally booked on Long March, including two for Intelsat, have since been rescheduled with other companies. □

Japan orbits satellite market

Tokyo. Japan will enter the commercial satellite launch business in four years' time, with a contract to launch ten satellites for a US company using a modified version of its H-2 rocket, according to Yasuhiro Moriguchi, director of the Science and Technology Agency's aeronautics and space development division.

Moriguchi said last week that Hughes Space and Communications International Incorporated will sign a contract for the launches next month with Rocket Systems Corporation (RSC), a consortium of 73 Japanese aerospace firms, banks and insurance companies set up several years ago to sell launches on the H-2 rocket. RSC refuses to confirm details of the deal, but admits that it is negotiating a 'lot purchase' with Hughes. □

Model shuttle glides home

Tokyo. The National Space Development Agency of Japan took a significant step towards the development of an unmanned space shuttle last Saturday (6 July) with the successful test-landing of its unmanned Automatic Landing Flight Experiment (ALFLEX) aircraft in Australia (above).

The 6.1-metre ALFLEX is a scale model of Japan's proposed unmanned space shuttle, HOPE, which is expected to be launched early in the next century. It was dropped by a helicopter from an altitude of 1,500 metres, and glided automatically to a runway 2.7 km away, using its own satellite navigation and flight control system. □



NASDA

A column from the HisTrap™ Kit, magnified 280%.

Would you like to save one hour in histidine-tagged protein purification?
To load, purify and elute in 15 minutes? Now it's possible—time after time.

HisTrap kit: saving an hour couldn't be simpler

The HisTrap™ Kit presents you with a solution so simple it works with a syringe. Inside the kit is all you'll need for up to ten affinity purification runs: pre-packed cartridges, ready-made stock solutions, a protocol that's easy to follow, and, of course, a syringe.

You'll get high purity recombinant proteins as the kit lets you use the well-known binding of poly(histidine) sequences to immobilized nickel. What's more, it saves you the effort of weighing out buffer salts.

Molecular biologists have been purifying recombinant proteins with Pharmacia Biotech media from the very beginning. Call us at 1 (800) 526 3593 from the USA, +81 (0)3 3492 6949 from Japan or +46 (0)18 16 50 11 from Europe and the rest of the world (or meet us on the Internet at <http://www.biotech.pharmacia.se>). Ask about the kit that saves an hour.

Pharmacia Biotech
Uppsala, Sweden. (And the rest of the world)



"Purity, simplicity and time savings: a product that gives you all these, lets you take a break from the lab for a bagel, coffee and...hmm, maybe a haircut?" Kyungsun Suh, Ph.D., at the Rockefeller University, New York, NY

Patents versus transplants

SIR — In 1989, Gluckman and colleagues reported that haematopoietic stem cells from umbilical-cord blood of a neonate could be used as an alternative to bone marrow to engraft the donor's HLA-identical sibling who had Fanconi's anaemia. Seven years later, this patient is alive and apparently cured. Since 1989, it has become clear that cord blood can be used to treat patients in need of a bone-marrow transplantation.

The use of cord-blood cells has several advantages, including a higher concentration of immature progenitors leading to better engraftment and reduced immune reactivity. This allows successful transplants in the absence of a donor matched at all 6 HLA loci. More than 200 allogeneic cord-blood transplants have been carried out, and autologous cord-blood cells have been used for gene therapy. Programmes of cord-blood banking have been established for family and unrelated transplants worldwide.

The technical, scientific, clinical and ethical aspects of the use of cord-blood stem-cells have been considered at a number of national and international meetings, and papers have been published in the scientific literature. These topics are the subject of continuing attention and evaluation by standing committees of the World Marrow Donor Association, the Immunobiology Working Party of the European Group for Blood and Marrow Transplantation, the Eurocord Transplant group and similar programmes in the United States.

It has recently been drawn to our attention that the clinical use of cord-blood stem cells, for all purposes, is the subject of a patent which is being actively disputed in the United States, which has been approved in Europe and which is being applied for in Japan. The holders of these patents presumably intend to exploit the use of cord-blood stem cells for commercial purposes and to oppose their use by transplanters unwilling to pay the patent fees. If these patents are enforced, the use of cord-blood cells for transplantation will not conform with the resolution of the International Society of Transplantation which establishes the basic principle that no part of the human body should be commercialized and that donation of organs or cells should be free and anonymous. For this reason, Eurocord has decided to oppose this patent.

The signatories of this letter believe that all procedures involving the transplantation of human organs, including haematopoietic stem cells, should be carried out only in an orthodox clinical setting where commercial considerations do not apply. Blood and organs are freely donated and there is an implicit agreement within the health-care community that they should not be used for the benefit of financial speculators. We con-

demn any attempt to patent a nonpharmacological method of using this or any other haematopoietic stem-cell source for treating patients with haematological diseases and we recommend that clinicians and scientists dissociate themselves from patents of this type, whether they have already been granted or are only in the application stage.

E. Gluckman*

Hôpital St Louis,

1 Avenue Claude Vellefaux,

75010 Paris, France

e-mail: eliane.gluckman@chu-stlouis.fr

R. O'Reilly, Bone Marrow Transplant Service, Memorial Sloan-Kettering Cancer Center, New York **J. Wagner**, Pediatric Bone Marrow Transplant Program, University of Minnesota, Minneapolis; **P. Rubinstein**, New York Cord Blood Bank

*On behalf of J. Goldman, J. J. Van Rood, J. Dausset & A. Gratwohl of the board of the European Blood and Marrow Society, the Eurocord transplant group, France Greffe de Moelle & the board of the French Society of Bone Marrow Transplantation.

Passive voice

SIR — I am an academically trained linguist (PhD, University of Chicago). I have spent 14 years teaching English composition and linguistics in various universities, where I have taught graduate seminars on the structure and process of written language, and 15 more as a professional writer. I am not a regular reader of your publication; my brother, who is, showed me Simon Leather's letter, which purports to make "the case for the passive voice" (*Nature* 381, 467; 1996).

About all that I can say with certainty regarding the passive voice is that it omits the performer of the action, for reasons that may be contextual (the reader already knows who or what performed the action, so mentioning the agent is redundant); rhetorical (either the agent is unknown or irrelevant, or the writer wishes to conceal his/her/its identity); structural (the writer wishes to keep sentence-topic consistent from one sentence to the next); or cultural (as in scientific writing, where use of the passive, rigidly enforced by senior members of the community, serves as a sign of in-group membership).

Leather goes far beyond these simple truths, with absolutely no scientific data. Where are the quantitative stylistic analyses, where is the behavioural or psycholinguistic research to support his statements that "[t]he use of the passive voice encourages disciplined writing" (indeed, "tenses must be used correctly" in all writing); that "using the active voice encourages possessiveness in the results and/or work;" that "the active voice [has a] tendency to foster colloquialisms;" that "the use of the active voice...

leads to an unwillingness to see [the author's] results contradicted" and to "the fabrication of results"? I know of no such research. I strongly doubt that it exists.

I consider Leather's letter an outrageous display of scientific hypocrisy. He makes dogmatic pronouncements on a subject he knows nothing about. It's as if I, armed only with my tenth-grade biology course, were to undertake to prove the superiority of mammals over reptiles, on the grounds that the former are cuddly and smart, whereas the latter are slimy and stupid.

Leather should practise what he preaches: of the 18 transitive sentences in his letter, only four were in the passive.

Alan M. Perlman

Kraft Foods, Inc.,

Northfield, Illinois 60093, USA

Public confidence

SIR — Eugene Wong opens his Commentary article, "An economic case for basic research" (*Nature* 381, 187–188; 1996), with two statements that he puts forward as uncontested and deeply troubling facts.

It is simply not true, in the United States at least, that basic research is being pressured downwards in favour of more applied work. In fact, the sitting US Congress is strongly on record as favouring basic research. Second, Wong cites no evidence for his stated "change in the public perception of the benefits of basic research". The evidence of which I am aware, drawn from many years of National Science Board public opinion polls (*Science & Engineering Indicators* — 1996) and current-year surveys in California, Florida and Texas commissioned by Research!America, tells the opposite story. For many years, at least 70 per cent of Americans of 18 years of age or older have said that they agree that "even if it brings no immediate benefits, basic science research which advances the frontiers of knowledge is necessary and should be supported by the Federal Government".

It is time for the research community to stop building a case based on overcoming presumed negatives and instead to build on the overwhelmingly positive regard in which the nonscientific public holds research, researchers and the institutions that house research.

Our rallying cry should be our pride in working in the public's interest; our strategies should include public outreach energized by a commitment to accessibility and accountability to a public that maintains a strong conviction that basic research is an essential investment in everyone's future.

Mary Woolley

(President)

Research!America,

1522 King Street,

Alexandria, Virginia 22314, USA

Mad cows, bats and baby milk

John Ashby

The intrinsic conflict between the mandated need to inform the public of potential health hazards and the need of the media for sensational headlines is threatening to compromise the scientific process.

PRELIMINARY findings of potential significance to human health are increasingly used by the media to cause alarm and controversy. After a day or two it is usually concluded that politically sensitive data are being withheld by incompetent authorities. In the United Kingdom, there is at present no shortage of examples — bovine spongiform encephalopathy (BSE) in cattle, phthalate esters in baby-milk formulae and the appearance of a rabies-infected bat.

Common to these examples is the publication of preliminary research findings. Two separate problems are evident. The first is that the media, in general, have no interest in equivocal conclusions based on preliminary data. They either ignore such reports, or transform them into firm indications of a problem. The second problem is that fear of the media charge of suppression of data is leading to the dissemination of ever-more preliminary data. These two trends feed on each other with potentially serious consequences for the scientific process.

Because conclusions can never be absolutely certain in a developing scientific study, communication to the media of interim conclusions relating to a matter of public health seems doomed to misinterpretation and subsequent confusion. For example, the media and the general public do not know how to respond to the hypothetical possibility that prion proteins may be transferred from BSE-infected cows to humans. Preliminary discussions of this hypothesis therefore rapidly degenerated into calls to name the cow, or to name the human — such certainty being the required prelude to the apportioning of blame. A related danger is that, as the media pressure mounts, cautious preliminary conclusions may be defended by scientists not directly involved. This can lead both to the transformation of preliminary conclusions into confirmed conclusions and to the neglect of alternative hypotheses (as is suggested¹ to have happened with Purdey's speculative chemical explanation for BSE).

Unfortunately, this developing 'yob culture' in media handling of scientific data could compromise the processes of data generation and dissemination. The main components of the scientific process are as follows. First, hypothesis, thesis and fact form a hierarchy, the steps between them being scaled by observations. When observations undermine an hypothesis, they

should be published by the media with the same enthusiasm as accompanies the announcement of confirmatory results. In practice, failed hypotheses usually die a lingering and obscure death. Second, observations should be confirmed before they are published. In matters of public health, however, preliminary results that would probably be unacceptable for publication by a scientific journal are often mandated to be communicated. Third, confirmed observations may be open to several possible interpretations, leading to the need for systematic studies of the sources of their variability. These often take some time to conduct, and their necessary absence from preliminary communications lies at the source of most health scares. Fourth, appropriate control observations should accompany each step of a scientific study in order to eliminate artefacts — again, such controls are often absent when preliminary results are announced. The neglect of these basic principles of good scientific practice can transform preliminary findings into a sequence of passing health scares.

The most recent example of this trend was the announcement that a single bat, probably blown across the Channel from France, was harbouring a variant of the rabies virus when it bit two women in England. The qualifications in that statement were made clear from the start, and essentially removed the incident from the realm of a public health alert. Experts interviewed on a domestic radio station a few hours after the initial announcement said that there was no need for general concern, and that the primary test result might have been caused by an interfering virus. These warnings should have been sufficient to ensure that the official statement was seen as the communication of preliminary data to an informed public by diligent health authorities. In fact, the media handled the incident as if a major outbreak of rabies had occurred.

The potential hazard posed to babies by low-level contamination of infant milk formulae with phthalate esters was recently recognized by the UK Ministry of Agriculture, Fisheries and Food (MAFF). A detailed exploratory study was mounted involving the collection of 59 samples of formulae and their analysis for the presence of 12 phthalate esters. The results of the initial survey were communicated to around 3,000 interested parties through the

MAFF *Food Safety Bulletin* and its more detailed surveillance information sheet². It was concluded that "these initial surveillance data and the limited biological studies available so far do not suggest that the concentrations of phthalates found in infant formulae would have any effect in humans". This conclusion was derived in the light of the cautious preliminary observation by R. M. Sharpe *et al.*³ that a small effect on testes size and sperm production was produced in young rats exposed to one of these phthalates *in utero*. The surveillance document went on to recommend steps to reduce further, *as a matter of principle*, phthalate levels in formulae, and follow-on studies to confirm and extend the results of the initial survey. The media reaction to this preliminary report was that a problem did in fact exist and that the most affected brands of formulae should be identified so that consumers could boycott them. Being a pilot study, the MAFF investigation had not studied batch-to-batch variations within each brand, so the questions posed by the media could not be answered. Yet again, the media were unable to respond in a balanced way to preliminary data that in fact pointed to the safety of a product.

Given the perceived need for the communication of preliminary data on matters of public health, the scientific process cannot be strictly adhered to by those responsible for the generation and communication of such data. This then leads to the vicious circle of an implication by the media of either incompetence or cover-up. So there seems to be an intrinsic conflict between the mandated need to inform the public of potential health hazards and the need of the media for headlines. The solution lies with general education in the scientific process and concerted efforts to inform the public about relative risks. Curiously, it is the tobacco industry that has taken the initiative in the latter, observing that the hazard of passive cigarette smoking equates to consuming three hamburgers a day. □

John Ashby is at Zeneca Central Toxicology Laboratory, Alderley Park, Cheshire SK10 4TJ, UK.

1. Booker, C. & North, R. *Daily Telegraph* (14 June 1996).
2. Food Safety Directorate *Food Surveillance Information Sheet*, March 1996 (MAFF, London 1996).
3. Sharpe, R. M. *et al.* *Envir. Hlth Perspect.* **103**, 1136–1143 (1995).

Breathing room for early animals

Andrew H. Knoll

Did an increase in atmospheric oxygen concentration some time during the Neoproterozoic era, 1,000 to 543 million years ago, make possible the diversification of multicellular animals?

THE fossil record shows a dramatic radiation in animal forms in the latest Proterozoic and early Cambrian epochs, the so-called Cambrian Explosion. But molecular clock^{1,2} and developmental³ arguments instead favour much earlier splitting of invertebrate lineages. These differing results could be reconciled if some environmental change allowed a sudden increase in body size, making animals (and their fossils) visible and opening up new morphological and behavioural avenues for diversification. Early advocates of the oxygen-enrichment hypothesis noted that macroscopic animals need more oxygen than their probable microscopic ancestors. For oxygen to be viewed as a direct influence on animal radiation, however, one must demonstrate not only that such animals require high oxygen concentrations, but that the critical environmental threshold (10% of the present-day atmospheric level, or higher) was crossed late in the Proterozoic aeon.

The first direct geochemical support for Neoproterozoic oxygen enrichment came from carbon isotopes, which show that the organic carbon burial rate increased strongly during the later Neoproterozoic era⁴⁻⁶. (Photosynthesis produces organic matter and oxygen, and aerobic respiration consumes both, so the potential for atmospheric oxygen enrichment arises when organic matter is shielded from oxidation by burial in sediments.) Unfortunately, carbon isotopes cannot provide a quantitative estimate of late Proterozoic oxygen concentration. On page 127 of this issue⁷, Donald Canfield and Andreas Teske approach the problem from a new perspective, presenting innovative arguments based on the biology and biogeochemistry of sulphur-metabolizing bacteria. If correct, they not only confirm the qualitative inference of Neoproterozoic oxygen enrichment, but indicate that environments capable of supporting large animals first stabilized at that time.

Bacterial reduction produces sulphates depleted in ³⁴S by approximately 20±10‰ (parts per thousand), yet sulphates and pyrites in Phanerozoic sedimentary rocks (less than 543 Myr old) commonly differ by as much as 50 to 70‰. Canfield and Thamdrup⁸ showed

experimentally that the additional fractionation needed to explain the sedimentary record can be accounted for by the oxidation of H₂S to sulphur, followed by bacterial disproportionation of the sulphur to sulphide and sulphate. H₂S generated in this way is depleted in ³⁴S by about 8‰, so the geochemical record can be explained by repeated cycles of oxidation



Late Neoproterozoic sedimentary rocks exposed in the glaciated highlands of Spitzbergen. These rocks preserve the fossils that document early animal evolution, as well as a sedimentary and geochemical record of coeval environmental change.

and disproportionation before the sulphide was sequestered as pyrite.

Canfield and Teske⁷ here show that the large isotopic fractionation observed in Phanerozoic rocks is not matched in the Proterozoic record. Fractionation in early Neoproterozoic and older rocks can be accounted for by sulphate reduction alone, leading Canfield and Teske to conclude that, before the Neoproterozoic era, oxygen levels in the ocean were too low to drive the repeated oxidation of H₂S inferred for younger periods.

Precisely when the transition occurred is difficult to determine from the limited data currently available. Sedimentary sulphides from the marine sandstones and shales of the Windermere Supergroup, Canada⁹, suggest that the game was over by about 650 Myr ago, and two data points from pyrites between about 800 and 850 Myr old⁷ hint at earlier oxygen enrichment. It may be overly simplistic to assume that oxygen concentration increased monotonically; sulphur isotope data could reflect fluctuating oxygen levels before stabilization about 650 Myr ago, close to the time when the ¹³C/¹²C ratio reached its highest value for the past two billion years⁶. That indicates a high rate of organic carbon burial, which is probably the mechanism behind the oxygenation event.

Because such a large increase in fractionation should have influenced the sulphur-isotope composition of sea water, Neoproterozoic sulphates might preserve a complementary record of environmental transition. In fact, the proposed increase

is enough to account for the so-called 'Yudomski Event', a latest Proterozoic rise¹⁰ of about 10‰ in sulphate ³⁴S. Certainly younger than 780 Myr, the Yudomski shift may correlate with the events of 650 Myr ago inferred from the bulk of sulphide data and from carbon isotopes¹⁰.

Canfield and Teske's second argument rests on the phylogeny and physiology of sulphur-oxidizing bacteria. These organisms live along the gradient between oxic and sulphide-rich environments, a habitat that Canfield and Teske suggest came into

- Runnegar, B. J. *Geol. Soc. Austral.* **29**, 395-411 (1982).
- Doolittle, R. F., Feng, D.-F., Tsang, T., Cho, G. & Little, E. *Science* **271**, 470-477 (1996).
- Davidson, E. H., Peterson, K. J. & Cameron, R. A. *Science* **270**, 1319-1325 (1995).
- Knoll, A. H., Hayes, J. M., Kaufman, A. J., Swett, K. & Lambert, I. B. *Nature* **321**, 832-838 (1986).
- Des Marais, D. J., Strauss, H., Summons, R. E. & Hayes, J. M. *Nature* **359**, 605-609 (1992).
- Kaufman, A. J. & Knoll, A. H. *Precamb. Res.* **73**, 27-49 (1995).
- Canfield, D. E. & Teske, A. *Nature* **382**, 127-132 (1996).
- Canfield, D. E. & Thamdrup, B. *Science* **266**, 1973-1975 (1994).
- Ross, G. M., Bloch, J. D. & Krause, H. R. *Precamb. Res.* **73**, 71-99 (1995).

- Strauss, H. *Precamb. Res.* **63**, 225-246 (1993).
- Moran, N. A. *Proc. natn. Acad. Sci. U.S.A.* **93**, 2873-2878 (1996).
- Krueger, D. M., Gustafson, R. G. & Cavanaugh, C. M. *Biol. Bull.* **190**, 195-202 (1996).
- Cary, S. C. *Molec. mar. Biol. Biotech.* **3**, 121-130 (1994).
- Ochmann, H. & Wilson, A. C. *J. molec. Evol.* **26**, 74-86 (1987).
- Logan, G. A., Hayes, J. M., Hieshima, G. B. & Summons, R. E. *Nature* **376**, 53-56 (1995).
- Holland, H. D. & Beukes, N. J. *Am. J. Sci.* **290A**, 1-34 (1990).
- Hofmann, H. J., Narbonne, G. M. & Aitken, J. D. *Geology* **18**, 1119-1202 (1990).
- Knoll, A. H., Bambach, R. K., Canfield, D. E. & Grotzinger, J. P. *Science* (in the press).

being only when oxygen levels reached a threshold somewhere between 5 and 18% of present-day levels. Given the capacity of prokaryotes to adapt rapidly to novel environments, the authors reason that a 16S ribosomal RNA clock age for the divergence of sulphur-oxidizers should also indicate when oxygen concentrations crossed the indicated threshold. Some sulphur-oxidizing bacteria live as symbionts in bivalve molluscs, prompting Canfield and Teske to calibrate their clock using fossil clams. The estimated date of divergence is 750 ± 320 Myr, in broad but welcome agreement with the geochemical data.

Few topics raise the dander of molecular evolutionists as readily as molecular clocks (sometimes written 'clocks', with the inverted commas speaking volumes). This particular analysis is likely to receive close scrutiny because it relies on endosymbionts for clock calibration. Moran¹¹ has shown that when bacterial symbionts in animals are transferred from

one generation to the next via the host's eggs, the resulting population structure promotes rapid genetic sequence divergence. At least three and possibly all symbioses between sulphide-oxidizing bacteria and bivalves are perpetuated in this manner^{12,13}. Thus, equal amounts of sequence divergence in branches occupied by free-living and by symbiotic bacteria may not translate into equal elapsed time. On the other hand, clock calibration using free-living proteobacteria¹¹ also suggests a late Proterozoic origin for sulphur-oxidizers, as do 'universal' rates of bacterial sequence divergence¹⁴ (also complicated by reliance, at least in part, on endosymbionts for rate estimation).

To distinguish the present hypothesis from alternative theories of Proterozoic environmental change^{15,16}, geochemists must sample the stratigraphic record more systematically, with increased emphasis on stratigraphic and sedimentological controls. Comparative biological

inferences will be greatly improved by better physiological data tied to more complete and better-resolved phylogenies.

Curiously, if Canfield and Teske are correct, the barrier of low oxygen may have fallen too early to fully explain the timing of Ediacaran diversification, a precursor of the Cambrian Explosion. Simple centimetre-sized fossils of probable metazoans¹⁷ occur in rocks older than 600 Myr, but diverse Ediacaran body fossils, tracks, and trails appear about 570 Myr ago. In between, the world experienced climatic and biogeochemical fluctuations whose severity matches or exceeds those associated with the great mass extinctions seen in younger periods¹⁸. In the end, atmospheric oxygen enrichment may prove to be just one of several key environmental influences on early animal evolution. □

Andrew H. Knoll is at the Botanical Museum, Harvard University, Cambridge, Massachusetts 02138, USA.

RNA

Aptly named aptamers display their aptitude

James R. Williamson

ADVOCATES of the view that life started in a primordial 'RNA world' hypothesize that the functions executed in the modern cell by proteins would have originally been performed by RNAs. In the modern biological world, the main ligands for functionally important RNAs are proteins, and RNA-catalysed reactions are limited to simple transesterifications. The structure of a 40-nucleotide RNA capable of specifically binding ATP, reported by Jiang *et al.* on page 183 of this issue¹, provides structural

a recent proliferation of RNA sequences that have been identified by directed-evolution methods carried out *in vitro*. Pools of up to 10¹⁵ different RNA molecules containing stretches of random sequences are effectively selected for their ability to bind to a particular ligand, and after many rounds of selection the 'winners' that are identified are termed aptamers². *In vitro* evolution has produced many such aptamers that bind such diverse ligands as proteins, amino acids, nucleotides and vitamin cofactors, and recently solution structures for different aptamers bound to flavin mononucleotide cofactor, and the amino acids arginine and citrulline, have been solved^{3,4}.

The original *in vitro* selection for aptamers that bind to ATP resulted in the identification of a short sequence motif capable of binding ATP with micromolar affinity⁵. Our ability to predict RNA structure from sequence is largely limited to partial secondary structures, and analysis of

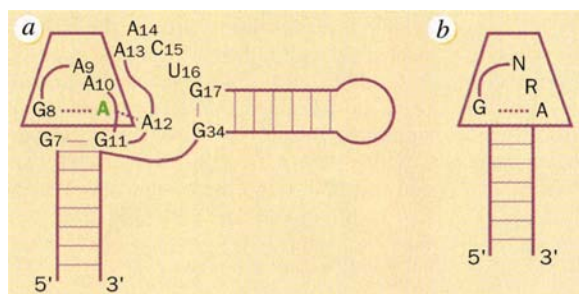
the ATP-aptamer sequences reveals an internal loop structure (shown in Fig. 1a of the paper on page 183), with an 11-nucleotide bulge on one strand opposite a one-nucleotide bulge on the other strand. At present, our understanding of RNA structure is insufficient to tell us how such

a secondary structure folds into a three-dimensional structure. That is why the solution structure of the complex of this aptamer with AMP reported by Jiang *et al.* is such a revelation.

The term aptamer comes from the Latin *aptus*, to fit, and the structure of the ATP-binding aptamer confirms that the name is indeed apt. The ATP 'fits' into a nucleotide-binding pocket made by an intricate network of hydrogen bonds to the adenine base and the ribose, and by stacking interactions where the bases lie flat upon one another. The interactions formed in the ATP-bound state are shown schematically in *a* of the figure shown here.

Inspection of this pocket reveals an old friend to those familiar with RNA structure, namely, the stable GNRA tetraloop (where G is guanine, N is any nucleotide, R is a purine and A is adenine). These loops are commonly found in large RNA molecules, presumably because they form a particularly stable structure and can thus serve as a cap for the ends of RNA helices⁶. The structure of the GNRA loop was one of the first RNA structures to be solved using nuclear magnetic resonance spectroscopy⁷, and the unusual stability of this particular loop sequence was attributed to the formation of a G-A base pair in the loop as well as stacking interactions. Part of the ATP-binding aptamer forms a structure essentially identical to the GNRA tetraloop, as shown in the accompanying figure. What is remarkable about the present structure is that the position of the A residue in the GNRA motif is occupied by the bound adenosine nucleotide.

The structure of the ATP-binding aptamer has also been solved by another group⁸, and this second paper includes some additional mutational data. The structure provides a role for all the



Comparison of the ATP-aptamer (a) and the GNRA tetraloop (b) structures. Important nucleotides in the ATP-aptamer are shown, with important hydrogen bonds indicated with dashed lines. The bound ATP is shown in green, and the GNRA motif is outlined in a box.

insights into how RNAs can bind to small molecules, as do their protein counterparts. The architecture of this ATP-binding RNA is quite remarkable, and may provide some clues to the nature of the denizens of the primordial RNA world.

The ATP-binding RNA is one example of

conserved bases in the ATP-binding aptamer motif. Interestingly, it is possible to substitute some of the conserved bases while maintaining the ability to bind ATP, albeit with slightly reduced affinity. The details of the hydrogen bonding between the G and A residues are slightly different in the GNRA loop and the ATP-binding aptamer, but the stacking, general arrangement of the nucleotides, and the backbone geometry, are quite similar. The GNRA motif is similar to the U-turns observed in the anticodon loop of transfer RNA and in the hammerhead ribozyme⁹. In addition, the bound adenine base is sandwiched directly between two nucleotides adjacent in the aptamer sequence, and this base-intercalation motif has also been observed in the tertiary core of tRNA structures. So even though our sampling of RNA structures to date is quite small, recurring structural themes are beginning to emerge.

Trying to understand how the current biology of RNA is related to the proposed RNA world amounts to molecular palaeontology: we try to reconstruct past events from the 'fossil record' of highly conserved RNAs such as ribosomal RNA or tRNA. Although we can endeavour to read the palimpsest¹⁰, the writing becomes more blurred the farther back in time we go. Although it is impossible to recapitulate hundreds of millions of years of evolution in the laboratory, the process of discovery of aptamer sequences can provide glimpses of possible alternative pasts. The ATP-binding aptamer presents a solution to the problem of how RNA might bind a nucleotide that is, in retrospect, all too obvious.

But does the ATP-binding aptamer resemble a primordial nucleotide-binding motif? I recall an artistic impression of *Tyrannosaurus rex* from the 1960s, where the claws and jaws were probably right, but the posture was wrong by today's standards, and the regal hide was portrayed in brilliant alternating bands of purple and orange. I suspect that the ATP-aptamer binding structure we now see has a similar relationship to a primordial nucleotide-binding motif — the bite is right, but the stripes are probably fanciful. □

James R. Williamson is in the Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

Prime time for neuropeptide Y

Rory Howlett

Two years ago, the identification¹ of the so-called 'satiety factor', leptin, caused enormous excitement in obesity research. Leptin is a hormone that circulates in the blood at levels paralleling those of fat reserves, and that binds specific receptors in the hypothalamus to mediate events in the brain's regulation of feeding behaviour. More recently, attention has switched to neuropeptide Y (NPY), the synthesis and release of which is inhibited as a consequence of leptin binding to its receptor. NPY itself is known to increase food intake when injected into the central nervous system.

Now come further results, from Gerald *et al.* (page 168 of this issue²), which should help clarify the role of NPY in the feeding pathway. But they also make all the more mysterious the finding of Erickson and colleagues³, reported in May, that mice in which the NPY gene has been knocked out maintain normal body weight.

Gerald *et al.* have cloned a new NPY-receptor subtype (Y5) from the hypothalamus of the rat, which could be the long-suspected NPY 'feeding receptor'⁴. A human homologue of Y5 has also been identified, and from circumstantial evidence it seems that it lies on the long arm of chromosome 4, closely linked but in opposite orientation to the gene encoding the Y1 receptor. The authors speculate that this locus might be associated with eating disorders or obesity in human populations.

The rat Y5 receptor is expressed in the brain, and to a lesser extent in the testes. What function it might have in reproduction is obscure, but the pattern of expression in the brain strongly suggests that it is an important regulator of feeding. The paraventricular nucleus (PVN) and the lateral hypothalamus, two areas already implicated in feeding, are both richly innervated by NPY-containing neurons; it now turns out that they express the Y5 gene transcript, as does the arcuate nucleus from which neurons run to the PVN. The authors reasonably conclude that it is likely that the Y5 receptor acts both presynaptically and postsynaptically, depending on where in the neuronal pathway it is expressed, perhaps modulating both the release of and response to NPY.

Intriguingly, the transcript is also present in thalamic nuclei from which neurons go both to the PVN and to the amygdala, an area often associated with emotional responses. Perhaps the tendency for some people to crave their favourite foods when under emotional strain should be blamed on this seemingly innocent piece of neuronal wiring.

The pharmacological profile of the Y5

receptor, as assessed in stably transfected cell lines, matches that of the Y1-like 'feeding receptor' predicted by previous studies. But more direct evidence for its involvement in feeding comes from experiments in which NPY-derived or related peptides, such as pancreatic peptide Y (PPY), are injected directly into the central nervous system of rats. Gerald *et al.* find that some peptides able to discriminate Y5 from other receptor subtypes can induce statistically significant increases in food intake within four hours of administration. The authors argue that these and other findings implicate the Y5-receptor subtype as the key mediator of the effects of NPY on feeding behaviour.

This is all very well, but there is a spanner in the works. In their recent paper, Erickson *et al.* found that mice deficient in NPY have normal food intake and body weight. Moreover, the mutant mice decreased their food intake and lost weight, initially to a greater extent than controls, when treated with recombinant leptin, suggesting that NPY is "not essential for certain feeding responses or leptin actions"³. As Stephens⁴ has said, it is possible that compensatory mechanisms can substitute for NPY in these mutant mice: the crossreaction of related peptides with Y-type receptors might be one such mechanism.

Even so, Gerald *et al.* don't hesitate to point out that the development of selective Y5 antagonists could lead to new treatments for obesity and related disorders. It is no coincidence that the authors are based at pharmaceutical companies. Effective anti-obesity drugs, particularly if free of untoward side effects, are likely to be highly profitable, and the incentives to invest in research and development are patent.

Whatever the outcome, one thing is sure: the synergy between molecular genetics and classical pharmacology may well transform our understanding of feeding regulation and its relation to general metabolic control. The two disciplines are as well matched as a rather older couple, immortalized in rhyme:

*Jack Sprat could eat no fat
His wife could eat no lean
And so between them both you see
They licked the platter clean.*

Rory Howlett is deputy biological sciences editor at Nature.

1. Jiang, F., Kumar, R. A., Jones, R. A. & Patel, D. J. *Nature* **382**, 183–186 (1996).
2. Ellington, A. D. & Szostak, J. W. *Nature* **346**, 818–822 (1990).
3. Patel, D. J. *et al.* *J. molec. Biol.* **258**, 480–500 (1996).
4. Yang, Y. *et al.* *Science* **272**, 1343–1347 (1996).
5. Sassanfar, M. & Szostak, J. W. *Nature* **364**, 550–553 (1993).
6. Woese, C. R., Winker, S. & Gutell, R. R. *Proc. natn. Acad. Sci. U.S.A.* **87**, 8467–8471 (1990).
7. Heus, H. & Pardi, A. *Science* **253**, 191–194 (1991).
8. Dieckmann, T. *et al.* *RNA* (in the press).
9. Jucker, F. M. & Pardi, A. *RNA* **1**, 219–222 (1995).
10. Benner, S. A., Ellington, A. D. & Tauer, A. *Proc. natn. Acad. Sci. U.S.A.* **86**, 7054–7058 (1989).

1. Zhang, Y. *et al.* *Nature* **372**, 425–432 (1994).
2. Gerald, C. *et al.* *Nature* **382**, 168–171 (1996).
3. Erickson, J. C. *et al.* *Nature* **381**, 415–418 (1996).
4. Stephens, T. W. *News and Views Nature* **381**, 377–378 (1996).

Filaments in creation's heart

John Bally

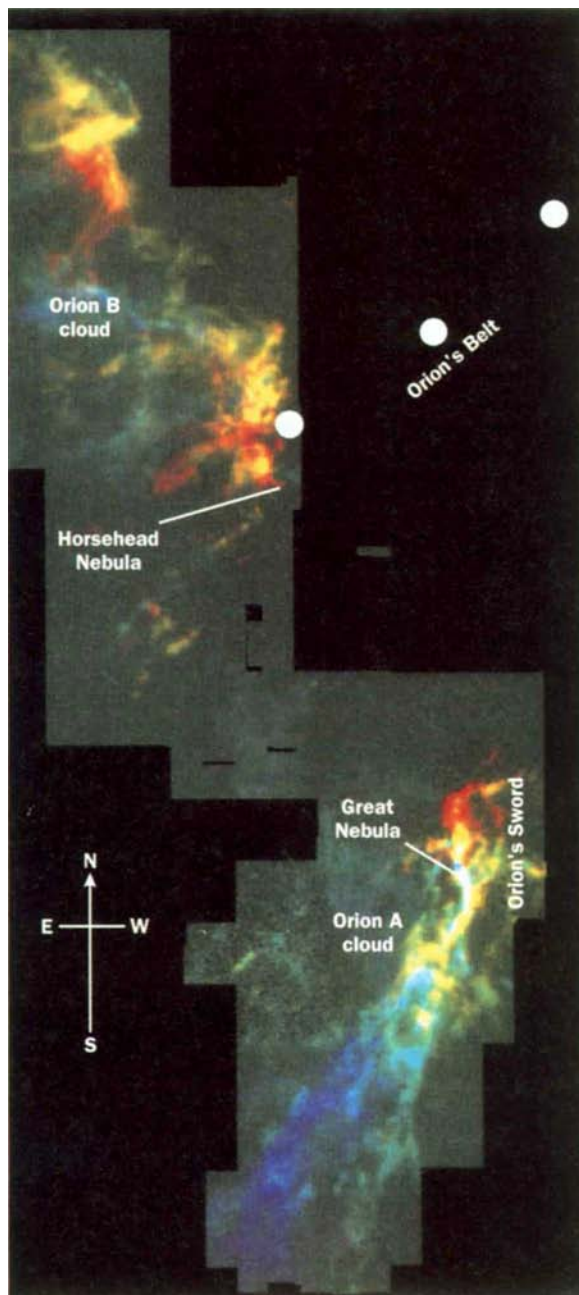
STAR formation is as fundamental to astrophysics as the formation of molecules is to chemistry. The star-formation process regulates the evolution of galaxies, the rate at which matter is locked into long-lived, low-mass stars, the chemical enrichment of the Universe by elements synthesized in the cores of massive stars, and, through the prodigious release of energy by these stars, the energetics of interstellar gas. Furthermore, planet formation is part of the same process. The constellation of Orion contains the best-studied star-forming region in the sky, and continues to yield new insights as ever more powerful telescopes are focused on it. On page 139 of this issue¹, Wiseman and Ho present a mosaic image of a remarkable system of dense filaments of molecular gas located directly behind the Great Nebula in Orion, the nearest region of massive star formation.

Some of these structures may be forming stars, others may still be in a pre-collapse phase. Why is the raw material of star formation so filamentary? Is this morphology a result of environmental factors — anisotropic compression by an external source, for example — or an inevitable feature of the gravitational collapse of molecular clouds? The answer may hold clues to the most basic aspects of star formation, such as the origin of the stellar mass spectrum and the tendency of young stars to be clustered.

Even in the absence of external forces, gravity can produce filamentary structures. If the initial distribution of cloud density is slightly anisotropic, the collapse time can be slightly shorter in one direction. A three-dimensional cloud then collapses into a sheet. The sheet can collapse further to form filaments, which then may condense into chains of condensations resembling 'beads on a string'. Simulations of the large-scale gravitational collapse of matter in the early Universe develop these structures. Magnetic fields, which are inherently anisotropic, may also contribute to the formation of pancakes, ropes and beads in an otherwise uniform medium.

Filaments can also be formed by external forces, such as the passage of shocks. When a planar shock wave hits a large

cloud, it will compress it into a dense sheet. Hydrodynamic instabilities or the anisotropic support of a magnetic field may lead to further collapse of a sheet to form filaments. Such stringy structure is



The two molecular clouds in Orion, seen in the 110-GHz emission line of carbon-13 monoxide. The filamentary structure that can be seen here extends to much smaller scales.

common in mature supernova remnants, where a shock wave is interacting with the interstellar gas.

The Great Nebula in Orion is a hot blister of photo-ionized gas, excited into luminescence by four very young, high-mass O-type stars called the Trapezium.

The Nebula marks the location of one of the highest concentrations of young stars in the sky. McCaughrean and Stauffer² have identified more than 700 stars less than a million years old within one parsec of the Trapezium. Furthermore, Hubble Space Telescope observations of young stars exposed by the harsh ultraviolet radiation of the Trapezium have provided compelling evidence for the existence of dense disks of circumstellar matter that may be the sites of present or future planet formation^{3,4}.

Star formation has been occurring in Orion for at least 12 million years, producing a sequence of stellar groups of decreasing age that trace the propagation of star formation from north of Orion's Belt towards the Great Nebula⁵. The Trapezium is the youngest group of massive stars in this sequence. The intense ionizing radiation, stellar winds and supernova explosions of dozens of these O- and B-type stars, have inflated an elongated 100- by 300-parsec 'superbubble' called Orion's Cloak, which extends from east of Orion to the constellation of Eridanus in the west⁶. Two giant molecular clouds have been overrun by the expansion of the bubble. Each cloud has a mass greater than 50,000 times that of the Sun, and they continue to produce more new stars east of Orion's Belt near the Horsehead Nebula and in Orion's Sword near the Great Nebula (see figure).

Infrared and radio sources, embedded jets, and outflows provide evidence for ongoing star formation in the high-density gas immediately behind the Great Nebula. At least one high-mass star with a luminosity of 10^5 Suns, known as *irc2*, is forming and driving fingers of shock-excited hydrogen emission into the molecular cloud⁷. The filamentary nature of this emission is probably produced by hydrodynamic instabilities in the outflow⁸.

On large (0.2–20-parsec) scales, most molecular clouds have filamentary structure. The figure shows an image of the Orion molecular clouds in 2.6-mm radiation, produced by the ground-state rotational transition of ^{13}CO , colour-coded according to the velocity of the emitting gas⁹. The southern molecular cloud in Orion points towards the centre of the Orion OB association like a dagger, suggesting that its cometary shape and the high density of its northern portion — where most of the star formation is taking place — may have been caused by the

expansion of the superbubble. Perhaps the filamentary structure is due to the combined forces of gravity and external compression?

Wiseman and Ho¹ have obtained more detailed images of the region (see Fig. 2 on page 140), using the Very Large Array (VLA) radio telescope to look at centimetre-wavelength transitions of ammonia. They show that the filamentary topology extends to scales of 0.02 to 1 parsec, at least an order of magnitude smaller than previously seen. On these smaller scales, the outflows produced by stars forming from the cloud, such as *irc2*, may be an important source of anisotropic compression contributing to the formation of the filaments, which in turn form more stars.

Further studies of the new filaments are needed to determine if they now contain stars, or are perhaps on the verge of forming new stars. Whatever the answer, the new observations provide tantalizing clues about the nature of hierarchical fragmentation in star-forming giant molecular

cloud cores, the process that leads to the formation of dense clusters of young stars. On the 200th anniversary of the publication of his nebular theory of the origin of the Solar System, Simon Laplace would be pleased that we are on the verge of seeing the process in a nearby cloud. □

John Bally is in the Astrophysical, Planetary and Atmospheric Sciences Department, University of Colorado, Boulder, Colorado 80309, USA.

1. Wiseman, J. J. & Ho, P. T. P. *Nature* **382**, 139–141 (1996).
2. McCaughrean, M. J. & Stauffer, J. R. *Astr. J.* **108**, 1382–1397 (1994).
3. O'Dell, C. R. & Wen, Z. *Astrophys. J.* **436**, 194–202 (1994).
4. McCaughrean, M. J. & O'Dell, C. R. *Astr. J.* **111**, 1977–1986 (1996).
5. Blaauw, A. in *The Physics of Star Formation and Early Stellar Evolution* (eds Lada, C. J. & Kylafis, N. D.) 125–154 (Kluwer, Dordrecht, 1991).
6. Cowie, L. L., Songaila, A. & York, D. G. *Astrophys. J.* **230**, 469–484 (1979).
7. Allen, D. A. & Burton, M. G. *Nature* **363**, 54–56 (1993).
8. Stone, J. M., Xu, J. & Mundy, L. G. *Nature* **377**, 315–317 (1995).
9. Bally, J., Stark, A. A., Wilson, R. W. & Langer, W. D. *Astrophys. J.* **312**, L45–L49 (1987).

TUMOUR SUPPRESSORS

Dispatches from *patched*

Ben-Zion Shilo

THE overlaps between gene pathways controlling embryonic development and those involved in regulating cell proliferation are becoming very compelling. A new celebrity that performs in both of these is the human *patched* gene, which plays many roles during development and now makes an entrance as a tumour suppressor. The evidence exposing this latest alias of *patched*, a faulty version of which is associated with a complex of cancers known as nevoid basal-cell carcinoma (NBCC) syndrome, was described last month in *Cell*¹ and *Science*².

Patched (*ptc*) was originally identified in the fruitfly *Drosophila* as a segment-polarity gene that is essential for embryonic patterning (not to be confused with the mouse *patch* mutation, in which the gene encoding a platelet-derived growth-factor receptor is deleted). It encodes a protein with 12 putative transmembrane domains, and may function as a receptor or a transporter^{3,4}. Mouse and chicken homologues of *ptc* have also been isolated^{5,6}.

The protein is a key element in the 'Hedgehog' signalling pathway, which is composed of an alternating series of inhibitory elements in both *Drosophila* and vertebrates⁷ (see figure). Ultimately, the secreted protein Hedgehog is responsible (directly or indirectly) for repressing the activity of the *ptc* protein, Patched. Patched normally acts through a pathway that is only partially characterized, re-

pressing transcription from downstream genes such as *ptc* itself, as well as preventing expression of members of the Wnt and Dpp/BMP families of morphogens, proteins that impart positional information. The introduction of Hedgehog thus relieves the inhibition conferred by Patched, allowing the downstream genes to be transcribed.

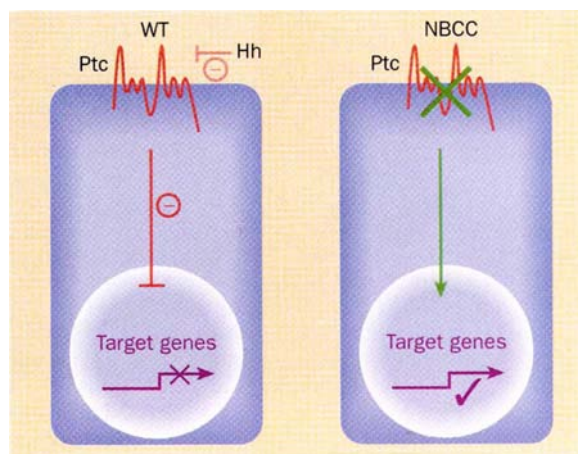
In *Drosophila*, the Hedgehog/Patched pathway is involved in a variety of processes, including segmentation in the embryo and the establishment of anterior-posterior polarity in the wing disc. The same pathway is also crucial during vertebrate development, inducing the formation of floor-plate and motor neurons by the notochord, and determining the anterior-posterior axis of developing limbs.

A growing number of genes have been found to play multiple roles during embryonic and post-embryonic development. Because an embryo will die early on if damaged copies of these genes are inherited from both parents (a homozygous mutation), it is not possible to study the effects of such mutations at later stages. To

circumvent the problem of these early abnormalities, the technique of mosaic analysis has been applied in organisms such as *Drosophila*. By generating clones of cells that are homozygous for the mutation of interest in the background of a heterozygous (and therefore viable, as only one of the gene pair is damaged) organism, it is possible to study the resulting effects (phenotypes) both within the clones and in the bordering normal tissues. Similar techniques are being developed in mice.

The demonstration that the NBCC syndrome (also termed basal-cell nevus carcinoma or Gorlin's syndrome) results from mutations in the human *ptc* gene provides a striking example of mosaic analysis in humans^{1,2,8}. The NBCC syndrome is an inherited, non-sex-linked (dominant autosomal) disorder characterized by predisposition to basal-cell carcinomas of the skin, medulloblastomas and ovarian fibromas, as well as a variety of malformations including pits of the palms and soles, rib and craniofacial alterations, and generalized overgrowth.

Haploinsufficiency for the NBCC locus on the one hand, and the time of sporadic loss of the second allele on the other, give rise to three sets of defects. First, symmetrical defects occurring in all patients, such as overgrowth, macrocephaly and facial dysmorphism, probably result from perturbation of dosage-sensitive pathways during development. Second, defects including malformed spine and ribs appear in multiple, random clusters within the same patient.



In normal, wild-type (WT) cells (left), Patched (Ptc), a 12-transmembrane-domain protein, represses the transcription of target genes. This inhibition is normally relieved only on introduction of the secreted protein Hedgehog. In NBCC cells, loss or inactivation of the second *ptc* allele is likely to result in constitutive induction of downstream genes, regardless of the presence or absence of Hedgehog.

The mosaic manifestation may be a consequence of early loss or inactivation of the normal *ptc* allele in a progenitor cell of the relevant tissue. The observed defects are consistent with the expression

sites and known functions of the Hedgehog/Patched pathway in development. Finally, loss of the normal *ptc* allele during adult life leads to multiple foci of basal-cell carcinoma.

It will be important to test whether homozygous loss of the *ptc* gene also occurs in the sporadic form of basal-cell carcinoma (BCC), which is the most common type of human cancer (750,000 estimated cases in the United States per year). Preliminary studies show that in 3 out of the 12 BCC tumours examined so far, the *ptc* region has been lost from one chromosome and there are mutations in *ptc* on the other^{1,2}.

The role of Patched as a tumour suppressor reveals new roles for the Hedgehog/Patched pathway, and suggests that Hedgehog is involved in the induction of cell proliferation in several adult tissues. Three conserved Hedgehog genes (*sonic*, *desert* and *indian*) have been identified in higher vertebrates, all of which are capable of activating the same pathway. It would be interesting to find out which of the three is expressed in skin. The common function of Patched in various developmental contexts and in different organisms is to suppress the expression of specific genes. Are these target genes similar during development and in the adult skin? Would the generation of small clones homozygous for a defective *ptc* gene also lead to tumour growth in *Drosophila*?

So *ptc* represents the first developmental gene that is also a tumour suppressor. Other developmental genes such as *Wnt1* and *Gli* (human glioblastoma gene, the homologue of the *Drosophila Ci* gene) were initially identified as oncogenes. In fact, both of these genes appear to be direct or indirect targets for repression by Patched, which may explain why *ptc* is a tumour-suppressor gene whereas they are oncogenes. In contrast to the Rb and p53 tumour-suppressor proteins which interact in the nucleus with components of the cell-cycle clock, Patched is a membrane protein regulating a complex cytoplasmic cascade terminating in transcriptional repression. Future research will tell us just how the Patched pathway impinges on the regulation of the cell cycle. □

Ben-Zion Shilo is in the Department of Molecular Genetics, Weizmann Institute of Science, 76100 Rehovot, Israel.

Turbulent excursions

Leo P. Kadanoff

A BREEZE swirls around with velocities that vary wildly in space and time — like many other physical systems, it is turbulent. Despite its familiarity and ubiquity, turbulence is ill understood. It presents two problems at once: determining flow velocities and figuring out how completely these velocities mix constituents suspended in the flow. But after decades of ambiguous results, two papers now bid fair to tell us the range of these turbulent fluctua-

space, a vortical motion develops, and these swirls quickly generate new vortices. Soon there is an avalanche of complexity (Fig. 1). This turbulent motion governs blood flow through arteries, weather patterns, air flow past rapidly moving vehicles and the chaos at the surface of the Sun.

Long ago we gave up the notion that we could make accurate, detailed, long-term predictions in turbulent situations.

Instead, we aim for the more modest goal of estimating the probability of the various kinds of events which might occur in a fluid. Applied work estimates the probability of interesting events — whether rain will fall tomorrow, for example. More fundamental studies concentrate on simpler events, such as the probability of observing a certain value of the difference in fluid velocity or contaminant concentration between two closely spaced points. For example, one can use an experimental picture of the mixing of two fluids (Fig. 2) to determine the probability of finding both fluids within a specified short distance.

The fundamental theory of turbulence was set forth in two papers by Kolmogorov published in 1941 and 1962 (and reviewed in ref. 2). Both of these describe the probabilities for observing the different possible values of $\Delta u(r)$, the velocity difference between two points separated by a distance r . The two Kolmogorov papers sketch out two quite different theories. The earlier one assumes a situation of 'simple scaling' in which the fluctuation in Δu has a 'typical' magnitude, predicted to vary as $r^{1/3}$. Events with differences far from this typical magnitude are then very improbable. In contrast,

the later paper, called K62, assumes no typical value of Δu ; rather, there is a tremendous range of event sizes. A situation in which some quantity varies so wildly that it has a range of characteristic scales is described by the word multiscaling. In the years since 1962, both points of view have had champions among experimenters, theorists and simulators.

Nianzheng Cao and colleagues make a computer simulation in a periodic box¹ and measure the probability distribution of sizes of Δu . They find far more variation than allowed by the K41 theory, and the number of outlier events can be fitted



FIG. 1 Water in turbulent motion. A jet flowing through a circular hole into a tank of still water is made visible using a fluorescent dye, illuminated by a thin sheet of laser light. This snapshot shows the complex patterns that result from the turbulent fluid motion. (From the laboratory of K. Sreenivasan, Yale University.)

tions. The computer calculation made by Nianzheng Cao, Shiyi Chen and Zhen-Su She¹ shows that the range of velocity fluctuations is too large to fit the 1941 Kolmogorov theory². For decades this theory was the standard of the field. Another apparent disagreement with the 1941 ideas appears in a theoretical and computational paper³ based on old theoretical work by Robert Kraichnan⁴. The new paper claims a determination of the probability distribution of the proportions of different constituents, which show anomalously large fluctuations.

If one puts a fluid into a situation in which its velocity varies rapidly enough in

1. Hahn, H. et al. *Cell* **86**, 841–851 (1996).

2. Johnson, R. L. et al. *Science* **272**, 1668–1671 (1996).

3. Hooper, J. E. & Scott, M. P. *Cell* **59**, 751–765 (1989).

4. Nakano, Y. et al. *Nature* **341**, 508–513 (1989).

5. Goodrich, L. V. et al. *Genes Dev.* **10**, 301–312 (1996).

6. Marigo, V. et al. *Development* **122**, 1225–1233 (1996).

7. Perrimon, N. *Cell* **80**, 517–520 (1995).

8. Levanat, S. et al. *Nature Genet.* **12**, 85–87 (1996).

quantitatively using a multiscaling theory. The practical implications of this result are that engineering design must take into account a huge range of possible values of every turbulent quantity. This wide variation can make all kinds of tasks



FIG. 2 A passive additive, coloured white, is mixed by the motion of a two-dimensional fluid flow. (From the laboratory of K. Sreenivasan, Yale University.)

harder, complicating the life of both the aeroplane designer and the scientist trying to simulate turbulent flow. On the other hand, the computation does provide support for some of the standard methods used to simplify numerical computations of turbulence, including large-eddy simulations.

Swirling fluids can also carry things around — other fluids or solids or heat, for example — which move and mix under the combined influence of the fluid motion and their own diffusion. Important issues, for example the extent of chemical reactions, depend on the degree of mixing of these passive constituents, and turbulence is crucial in obtaining good mixing. Recent calculations have concentrated on fluctuations in the spatial distribution of temperature differences, driven by rapid variations in the prescribed velocity field. The rapidity of the fluctuations makes the calculations possible.

Many authors agree that these temper-

ature differences also show the extreme variability of multiscaling^{5–7}. In fact, Kraichnan *et al.*³ claim that they have found an exact solution to the multiscaling problem that is robust under small changes in the way the flow is set up. They calculate the probabilities of outlier events, those with either exceptionally large or exceptionally small temperature differences. Their calculation implies that thermal mixing is, at some points in space and time, much less effective than one might expect on the basis of normal probability distributions.

If true, this would be a notable breakthrough indeed. The authors do point out possible weak points in their argument, but on balance they make a strong, almost compelling case for their result, especially

as their numerical simulations back up their main conclusions. Nevertheless, other authors^{5–7} look at the problem slightly differently, and fail to find the claimed results and/or robustness.

There is a real disagreement here. The difficult technical issues will not be settled in a day, but once they are settled we can expect a considerable advance in our understanding and control of mixing processes in swirly fluid flows, aiding the design of all kinds of devices which use turbulence, from aeroplanes to chemical-reaction vessels. □

Leo P. Kadanoff is at the James Franck Institute, University of Chicago, 5640 South Ellis Avenue, Chicago, Illinois 60637, USA. l-kadanoff@uchicago.edu

CIRCADIAN RHYTHMS

Ion channels get the message

Joseph S. Takahashi

OVER the past few years, we have learned a great deal about circadian clocks and how they function, both at the organismal and the molecular levels. We know that circadian pacemakers are discretely localized and can be isolated for study *in vitro*; that these pacemakers are cell autonomous; and that the fundamental molecular elements of the clock mechanism can and are being defined^{1–3}. In the fruit-fly *Drosophila*, the *period* and *timeless* genes encode elements of a transcription–translation loop oscillator involving negative autoregulatory feedback^{4–6}. An analogous mechanism holds for the *frequency* gene in the fungus *Neurospora*⁷. Nonetheless, we still know very little about how the circadian molecular machinery gets its message to the cell membrane and to other cells.

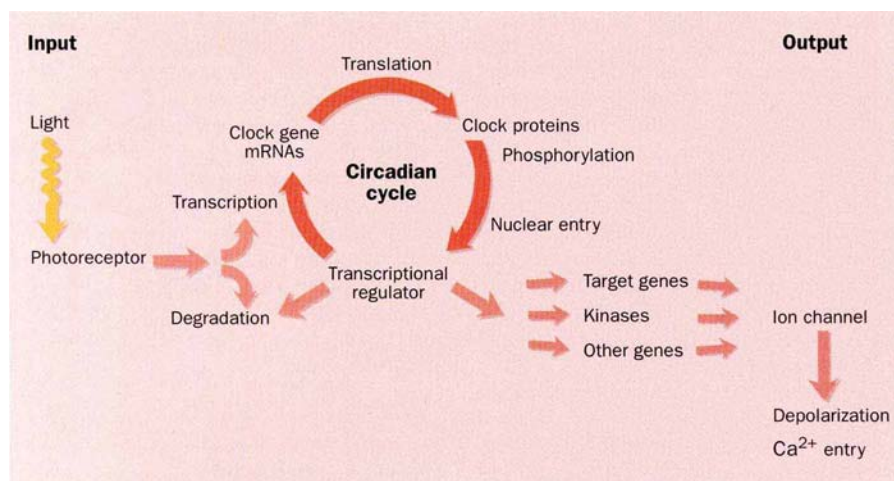
On page 165 of this issue⁸, D'Souza and Dryer report the discovery of a previously undescribed class of cationic channel, called I_{LOT} — for 'long open time' — that is under direct circadian control in chick pineal cells. The channel is nonselective for cations, has an intermediate unitary conductance (40 pS) with an unusually long open time, and does not appear to be gated by voltage or soluble second messengers. Most notably, D'Souza and Dryer find that the channel is active at night but not during the day in chick pineal cells maintained on light–dark cycles, and that the rhythm in channel activity persists in constant darkness *in vitro*. Previous work has suggested circadian regulation of a K^+ conductance in molluscan circadian pacemaker cells², but to my knowledge the experiments by D'Souza and Dryer are the first description, at the single-channel level, of an ionic channel that is under direct circadian regulation.

In vertebrates, circadian pacemakers are located in three structures in the central nervous system. These are the hypothalamic suprachiasmatic nucleus, the pineal gland and the retina. In each of these tissues, circadian rhythms can be expressed at the cellular level *in vitro*, suggesting that in multicellular organisms the circadian clock is a cellular entity and that an ensemble of cellular circadian oscillators act in concert to generate coherent pacemaker output at the tissue level.

Arguably, the chick pineal is one of the best cellular model systems for a vertebrate circadian oscillator⁹, because each cell (pinealocyte) contains all the elements of a circadian system — a photoreceptive entrainment pathway; a circadian oscillator; and an output pathway that regulates a rhythm in the biosynthesis of the now infamous molecule melatonin¹⁰, a hormone implicated in signalling diurnal and seasonal information in vertebrate organisms. Interestingly, chick pineal cells express a photopigment called pineal-opsin, as well as a number of retinal photoreceptor-specific proteins including a cyclic-GMP-activated channel^{11–13}.

The rhythm in melatonin synthesis is controlled at the level of an enzyme, arylalkylamine *N*-acetyltransferase (NAT), which is regulated in part by a circadian rhythm in abundance of messenger RNA¹⁴. In chick pineal cells, NAT is positively regulated by both cyclic AMP and the intracellular concentration of calcium ions, $[Ca^{2+}]_i$, which act synergistically. There is a circadian rhythm in cAMP levels which is generated endogenously in chick pineal cells, and evidence for an increase in $[Ca^{2+}]_i$ and Ca^{2+} /calmodulin-

1. Cao, N., Chen, S. & She, Z.-S. *Phys. Rev. Lett.* **76**, 3711–3714 (1996).
2. Frisch, U. *Turbulence* (Cambridge Univ. Press, 1995).
3. Kraichnan, R. H., Yakhot, V. & Chen, S. *Phys. Rev. Lett.* **75**, 240–244 (1995).
4. Kraichnan, R. H. *Phys. Fluids* **11**, 945–953 (1968).
5. Shraiman, B. I. & Siggia, E. D. *C.r. hebdom. Séanc. Acad. Sci., Paris Sér. IIb* **321**, 279–284 (1995).
6. Chertkov, M., Falkovich, G., Kolokolov, I. & Lebedev, V. *Phys. Rev. E* **52**, 4924–4941 (1995).
7. Gawedzki, A. & Kupianen, A. *Phys. Rev. Lett.* **75**, 3608–3611 (1995).



Possible control of the output of circadian rhythms by a transcription-translation loop oscillator — according to this proposal, output would be regulated by a circadian rhythm in expression of target genes such as those encoding kinases, ion channels and other regulatory elements.

stimulated adenylate cyclase activity at night. All of this suggests that the circadian clock machinery in chick pineal cells controls melatonin synthesis by regulating either cAMP levels or $[Ca^{2+}]_i$ on a circadian basis. How the chick pineal clock regulates these two second messengers remains an open question.

With D'Souza and Dryer's identification of I_{LOT} , one can imagine another scheme of events. Instead of directly regulating components of the cAMP or Ca^{2+} systems, the circadian clock would regulate the expression of I_{LOT} with a nocturnal profile (high at night, low in the day). Because I_{LOT} is a nonselective cation channel, pinealocytes would be more depolarized at night, and Ca^{2+} would enter the cell through both I_{LOT} and voltage-sensitive calcium channels. The elevation of $[Ca^{2+}]_i$ could then activate Ca^{2+} /calmodulin-stimulated adenylate cyclase (type I) and lead to an elevation in levels of intracellular cAMP. The joint increase in $[Ca^{2+}]_i$ and cAMP would then optimally stimulate NAT and melatonin synthesis at night. Such a rhythm of Ca^{2+} influx could provide one of the first examples of how the circadian oscillator mechanism can signal its output to the rest of the cell.

D'Souza and Dryer's results are of significance not only in providing a mechanism for circadian output in chick pineal cells in particular, but also in offering insight into a more general mechanism for other systems (especially neural circadian pacemakers such as the mammalian suprachiasmatic nucleus, retinal oscillators and invertebrate circadian systems). According to this view, the transcriptional regulation and expression of a channel on a circadian basis would provide an attractive mechanism for a transcription-translation loop oscillator to regulate the output of a neural circadian pacemaker cell by modulating either

membrane voltage or excitability (see figure).

Obviously, the discovery of I_{LOT} creates the immediate opportunity for characterizing it at the molecular level and cloning it. In addition, the mechanism by which the activity of I_{LOT} is regulated by the clock mechanism will be of much interest. Is I_{LOT} activity regulated primarily by phosphorylation, or by its level of expression? If I_{LOT} expression is regulated, does the circadian transcription-translation loop oscillator system (hypothesized to exist in vertebrates as well as *Drosophila* and *Neurospora*) control the transcription of the I_{LOT} gene as a consequence of the negative autoregulatory feedback loop, or are other steps interposed? Whatever the answers, they should be interesting — there's clearly I_{LOT} to do! □

Joseph S. Takahashi is in the NSF Center for Biological Timing, Department of Neurobiology and Physiology, Northwestern University, 2153 North Campus Drive, Evanston, Illinois 60208-3520, USA.

1. Takahashi, J. S. *A. Rev. Neurosci.* **18**, 531–553 (1995).
2. Michel, S., Geusz, M. E., Zaritsky, J. J. & Block, G. D. *Science* **259**, 239–241 (1993).
3. Welsh, D. K., Logothetis, D. E., Meister, M. & Reppert, S. M. *Neuron* **14**, 697–706 (1995).
4. Hardin, P. E., Hall, J. C. & Rosbash, M. *Nature* **343**, 536–540 (1990).
5. Zeng, H., Hardin, P. E. & Rosbash, M. *EMBO J.* **13**, 3590–3598 (1994).
6. Sehgal, A. et al. *Science* **270**, 808–810 (1995).
7. Aronson, B. D., Johnson, K. A., Loros, J. J. & Dunlap, J. C. *Science* **263**, 1578–1584 (1994).
8. D'Souza, T. & Dryer, S. E. *Nature* **382**, 165–167 (1996).
9. Takahashi, J. S., Murakami, N., Nikaido, S. S., Pratt, B. L. & Robertson, L. M. *Recent Prog. Horm. Res.* **45**, 279–352 (1989).
10. Turek, F. W. *Nature* **379**, 295–296 (1996).
11. Okano, T., Yoshizawa, T. & Fukuda, Y. *Nature* **270**, 94–97 (1994).
12. Max, M. et al. *Science* **267**, 1502–1506 (1995).
13. Dryer, S. E. & Henderson, D. *Nature* **353**, 756–758 (1991).
14. Klein, D. C., Roseboom, P. H. & Coon, S. L. *Trends Endocr. Metab.* **7**, 106–112 (1996).

DAEDALUS

Loaded crystal

THE traditional moderator for a nuclear reactor is graphite. Diamond, or even C_{60} , might be better, but are too expensive. But Daedalus likes the idea of exposing diamond to slow neutrons, for the free neutron is unstable. In a few minutes, it decays to a proton and an electron — a potential hydrogen atom. Now the diamond lattice is a web of carbon cages, each of three cyclohexane 'chairs' back to back, and enclosing a cavity about 0.5 nm across. A hydrogen atom would be a snug fit in such a cavity. So Daedalus reckons that in a flux of lethargically slow neutrons, a diamond would gradually fill up with hydrogen atoms. Each would be trapped in its carbon cage and unable to get at its neighbours.

The product, hydrogen-intercalated diamond (or 'Hidi'), would be a most intriguing substance. For a start, it would probably be the most rigid material known — superior even to diamond itself. The close-fitting hydrogen atoms would stiffen its lattice dramatically. Its refractive index would also exceed that of diamond. But as a gemstone it would have a serious drawback: it would be highly explosive. For atomic hydrogen is the most energetic substance known; its chemical energy is some 40 times that of TNT. If roughly handled, Hidi would explode to methane and amorphous carbon with appalling violence.

With luck, the trapped hydrogen atoms might be induced to combine peacefully. Gentle heat treatment might allow them gradually to 'hold hands' through their carbon cages, going to stable dihydrogen molecules. Dihydrogen-intercalated diamond, Dihidi, would be quite stable.

Hidi might also be stabilized by a magnetic field. The unpaired electron on each trapped hydrogen would give Hidi an unprecedented paramagnetic spin density. By aligning all their spins, the field would prevent the electrons from bonding. Indeed, they would generate a field strong enough to keep themselves aligned without further aid. So Hidi could be magnetized to a transparent and amazingly strong permanent magnet, still capable of exploding, but far more stable.

Hidi will be vastly expensive. But Daedalus recalls that hydrocarbon vapour passed over a suitable hot surface can deposit thin-film diamond. The film always contains hydrogen; indeed, atomic hydrogen is a crucial catalyst in the reaction. Could the process be tweaked to deposit Hidi? If so, it might become cheap enough to employ in magneto-optic modulators, micro-motors, thin-film inductors and magnetic limpet mines. Dihidi would be ideal for jewellery, lenses and ultra-high-strength reinforced laminates.

David Jones

Light on butterfly mating

SIR — The existence of extraocular photoreceptors is widespread within the animal kingdom¹. In arthropods such receptors exist in various parts of the body²⁻⁷, but their function is unknown in any species. Here we demonstrate that the light sense of genitalia in butterflies, which is mediated by two pairs of photoreceptor neurons (genital photoreceptors)^{8,9}, is crucial in achieving copulation.

The mating behaviour of the Japanese yellow swallowtail butterfly, *Papilio xuthus*, consists of six steps. The male first approaches a female (step 1), and gets into the 'venter-to-venter' position with the female, often opening the valva, the structure with which he supports her genitalia from the side during copulation (step 2). He then searches the female's genitalia by touching her abdomen with his own genitalia exposed between the fully-opened valva (step 3), firmly clasps her genitalia by using the superuncus and scaphium⁸. When the male has located the female's genitalia correctly (step 4), he inserts the penis, ejaculates, and then plugs the vagina with a sphragis by keeping the 'end-to-end' copulatory posture (step 5, Fig. 1), after which the pair finally separates (step 6).

Copulation requires precise alignment of the genitalia. How do butterflies confirm the alignment? What happens if the light sense of the genitalia is removed? We tested the effect of ablation of P1s, one of two pairs of genital photoreceptors, on mating behaviour. We bilaterally ablated P1s either by applying heat with a fine heat probe or by painting black material on the P1 site⁹.

Figure 2 summarizes our results: 66% of the intact males successfully copulate (copulation success rate = 66%) with intact vir-

gin females. However, the success rate falls to 28% when male P1s were bilaterally heat-ablated. The remaining 39% of P1 heat-ablated males fail at step 3. When P1s are painted black, the success rate is 23%, significantly smaller than that of either intact or control males where P1s are covered with clear paint. We found that 46% of the black-painted males stop mating behaviour at step 3. These results suggest that males require P1 input for precise location of the genitalia during copulation. P1 ablation in females does not affect the copulation success rate of the treated females with intact males.

It has been tempting to assume that the copulation is basically controlled by the mechanical sense of the genitalia¹⁰. The success rate of males whose mechanoreceptors on the scaphium are heat-ablated reduces to 44%, indicating the necessity of mechanical sense for achieving copulation. This may explain why one third of P1 heat-ablated males can still copulate: males locate females' genitalia by using the mechanical sense. We note that another pair of genital photoreceptors, the P2s (ref. 8), may also contribute to mating behaviour, although this possibility has not yet been tested.

How do P1s control mating behaviour? To address this question, we electrophysiologically recorded the P1 response from

mating males and revealed that the response sharply drops when the male correctly locates the female's genitalia. We therefore hypothesize that the drop in the P1 response informs the male that the female's vagina is correctly positioned for penis insertion. If the genitalia are not correctly aligned, the coupling would be incomplete, leaving some space through

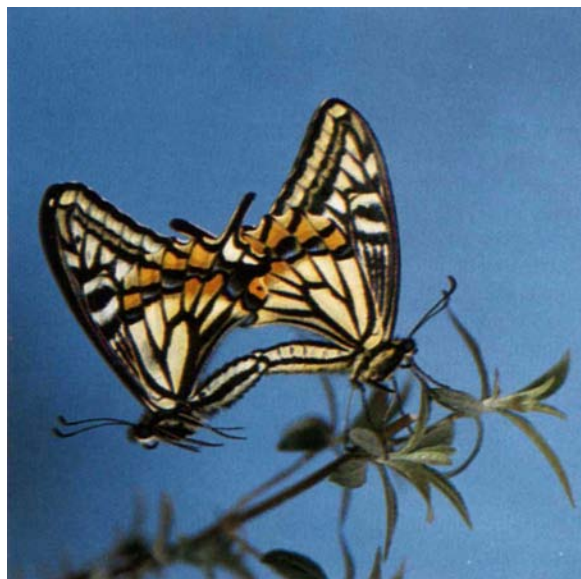
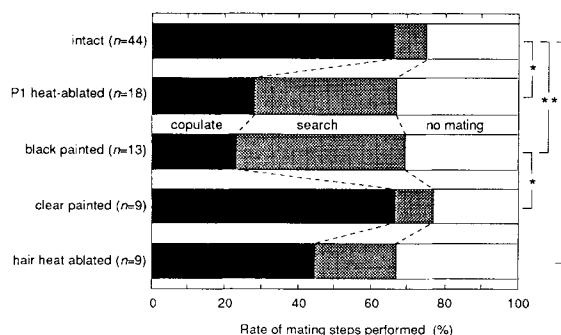


FIG. 1 A mating pair of *P. xuthus* in the copulatory posture (step 5). Up, female; Down, male. The male uses the light input from the genitalia to locate the female's genitalia upon copulation.

FIG. 2 Effect of P1 ablation on mating behaviour in males. Black bar, successfully copulated; hatched bar, stopped at step 3 (the search for female's genitalia); white bar, failed to perform mating behaviour. Male P1s were bilaterally ablated either by treating the P1 site with heat or by painting the whole scaphium with black mascara. The percentage of males that failed to perform mating behaviour (white bar) is not significantly different between treatments, indicating that there is no effect on the motivation for mating. We also ablated mechanoreceptive hairs on the scaphium, which was done by heating the receptor cell bodies from outside with a fine heat probe. The behavioural experiments were done in an outdoor cage (11 x 7 x 3 m). Ten to twelve intact virgin females were positioned in the cage by attaching the dorsal cuticle of the thorax with beeswax to the bottom end of an insect pin whose upper end was fixed to a horizontal bar set 2 m above the ground. We subsequently released 3-7 treated males. For each male, we recorded the behavioural step at which mating behaviour stopped. Mann-Whitney *U*-test was used for analysing the statistical difference in success rate between treatments (*, $P < 0.01$; **, $P < 0.05$).



which light can enter, so that the P1 response continues. In this case, we presume that the male releases the clasp and goes back to the search step. The males with P1 painted black never experience such a response drop as there is little P1 activity to begin with, but continue the genitalia search even when the female's genitalia are correctly aligned, until they finally give up.

K. Arikawa

D. Suyama

T. Fujii

Graduate School of Integrated Science,
Yokohama City University,
22-2 Seto, Kanazawa-ku,
Yokohama 236, Japan.

- Yoshida, M. in *Handbook of Sensory Physiology* (ed. Autrum, H.) 582-640 (Springer, Berlin, 1979).
- Ichikawa, T. *Zool. Sci.* **8**, 471-476 (1991).
- Kennedy, D. J. *gen. Physiol.* **46**, 551-572 (1963).
- Maitland, D. P., Heitler, W. J. & Laverack, M. S. *J. comp. Physiol. A* **168**, 513-520 (1991).
- Sandeman, D. C., Sandeman, R. E. & de Couet, H. G. *J. Neurobiol.* **21**, 619-629 (1990).
- Thompson, K. S. J. & Bacon, J. P. *J. comp. Physiol. A* **168**, 619-630 (1991).
- Yamashita, S. & Tateda, H. *J. comp. Physiol.* **150**, 467-472 (1983).
- Arikawa, K., Eguchi, E., Yoshida, A. & Aoki, K. *Nature* **288**, 700-702 (1980).
- Miyako, Y., Arikawa, K. & Eguchi, E. *J. comp. Neurol.* **327**, 458-468 (1993).
- Silberglied, R. E. in *The Biology of Butterflies* (eds Vane-Wright, R. I. & Ackery, P. R.) 207-223 (Academic, London, 1984).

Positive effects of pollution?

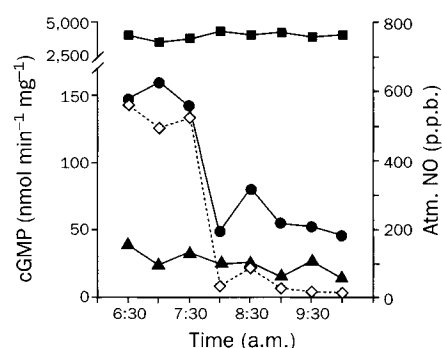
SIR — It took more than a decade to realize that the toxic gas nitric oxide (NO), a common air pollutant, is a signalling molecule in the body¹. As most of the actions of NO in the cardiovascular and nervous systems are mediated by soluble guanylyl cyclase², this enzyme represents the main intracellular receptor for NO. High-affinity binding of NO to the prosthetic haem group of soluble guanylyl cyclase results in a marked increase in cyclic (c)GMP formation³. Here, we demonstrate that soluble guanylyl cyclase purified from bovine lung is highly sensitive to small amounts of atmospheric NO generated by traffic and industrial combustion.

We measured the activity of this enzyme in close proximity to the Berlin inner city circular highway every half hour (see figure), and found that it paralleled nearby atmospheric NO concentrations (correlation coefficient 0.99), yielding as much as sixfold activation at an NO concentration of 550 parts per billion (p.p.b.).

We found further evidence for the stimulatory role of atmospheric NO on soluble guanylyl cyclase by showing that enzyme activity is not affected by atmospheric NO concentration in the presence of the NO scavenger oxyhaemoglobin (see figure). Maximally stimulated guanylyl cyclase activity (10 μ M diethylamine-NO complex) is constant over the indicated time and is independent of the NO concentration in the atmosphere. A plot of atmospheric NO concentration versus guanylyl cyclase activity in three independent experiments on different days revealed a linear relationship (y-intercept: 44, slope: 0.2) with a correlation coefficient of 0.94. Other air pollutants such as NO₂, CO, SO₂ and O₃ have correlation coefficients of 0.42, 0.87, 0.80 and -0.54, respectively.

Of these gases, only CO has been reported to stimulate soluble guanylyl cyclase⁴; but in contrast to the 400-fold stimulation by NO^{5,6}, CO stimulates the enzyme by a factor of only 4–6. Because of its higher concentration and much higher efficiency, we conclude that indeed NO,

Effect of atmospheric NO concentrations on the activity of purified bovine soluble guanylyl cyclase. Guanylyl cyclase activity was measured every half hour 30 feet from the Berlin city circular highway during the indicated time range in the absence of oxyhaemoglobin (filled circles) or in the presence of oxyhaemoglobin (65 μ M; filled triangles) or the NO-liberating diethylamine-NO complex (10 μ M; filled squares). The NO concentration in the atmosphere (open diamonds) was measured by chemiluminescence by the Department for City Development and Environment of Berlin at the same site. NO values are means of two consecutive 3-min measurements (18 averaged 10-s measurements) before and after the full or half hour. The marked decrease in atmospheric NO with concomitant reduction of cGMP production around 7:30 a.m. was due to sudden onset of wind.



rather than CO, is the primary stimulant of guanylyl cyclase in polluted air.

As soluble guanylyl cyclase is highly abundant in lung tissue, these results suggest a physiological effect of atmospheric NO in the lung. Indeed, beneficial consequences of inhaled NO (especially at concentrations of 100 to 1,000 p.p.b.^{7,8}) have been demonstrated in patients suffering from pulmonary hypertension and severe acute respiratory failure⁹. These atmospheric NO concentrations lead to 4–10 fold activation of guanylyl cyclase purified from lung. Thus, the data presented here argue in favour of using inhalation of low

concentrations of NO to produce a cGMP-mediated vasodilatory effect in treating lung disorders. There seem to be no effects on blood haemoglobin of inhaled NO^{7,8}, although there is recent evidence¹⁰ that haemoglobin is a transport molecule for NO.

**Andreas Friebe
Jürgen Malkewitz
Günter Schultz
Doris Koesling**

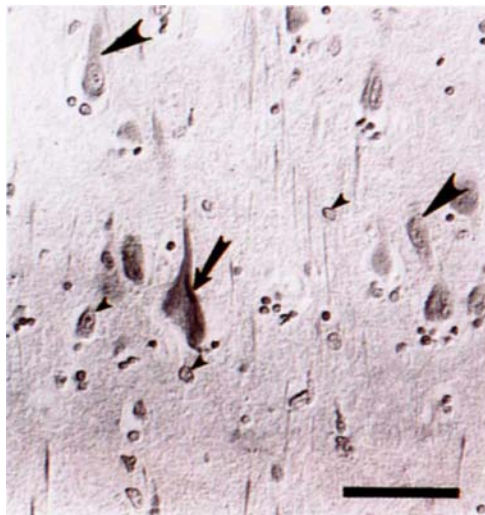
*Institut für Pharmakologie,
Freie Universität Berlin,
D-14195 Berlin,
Germany*

Oxidative damage in Alzheimer's

SIR — Oxidative stress results in the formation of free carbonyls derived from lipids, proteins, carbohydrates and nucleic acids¹. For example, metal-catalysed 'site-specific' oxidation of several amino-acid

side-chains produces aldehydes or ketones², and peroxidation of lipids generates reactive aldehydes such as malondialdehyde and hydroxynonenal. Here, we describe a highly sensitive, specific immunocytochemical technique, by using *in situ* DNP labelling linked to an antibody system, specifically to detect biomacromolecule-bound carbonyl-reactivity. This technique allows exact determination of the subcellular location of oxidative damage. We applied this method to tissues from cases of Alzheimer's disease, because oxidative events including lipoperoxidative, glycoxidative and other oxidative protein modifications have been reported and implicated in the pathogenesis of this disease³.

We studied non-aldehyde fixed (methacarn) hippocampal brain tissue from six Alzheimer's cases (age 76–90), three age-matched controls (age 70–78) and three younger controls (age 32–64). We sequentially rehydrated tissue sections, incubated them for 20 min in 3% H₂O₂ to inhibit endogenous peroxidase activity, and exposed them for 1 hour to 0.01% 2,4-dinitrophenylhydrazine (DNP). We detected bound DNP immunocytochemically with a rat monoclonal antibody against DNP followed by peroxidase-linked goat



Cases of Alzheimer's disease showing free carbonyls in neurofibrillary tangles (arrows) and in cytoplasm of neurons (large arrowheads) and nuclei of both glia and neurons (small arrowheads). By contrast, in age-matched and young controls, as well as for Alzheimer's samples after reduction of carbonyls by sodium borohydride (NaBH₄), none of these structures is labelled. Scale bar, 50 μ m. Further details of methods are available directly from M.A.S. and will be published in full elsewhere.

- Palmer, R. M. J., Ferridge, A. G. & Moncada, S. *Nature* **327**, 524–526 (1987).
- Ignarro, L. J. *Biochem. Pharmacol.* **41**, 485–490 (1991).
- Arnold, W. P., Mittal, C. K., Katsuki, S. & Murad, F. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3203–3207 (1977).
- Brüne, B., Schmidt, K.-U. & Ullrich, V. *Eur. J. Biochemistry* **192**, 683–686 (1990).
- Humbert, P. et al. *Eur. J. Biochemistry* **190**, 273–278 (1990).
- Stone, J. R. & Marletta, M. A. *Biochem. J.* **35**, 1093–1099 (1996).
- Gerlach, H., Pappert, D., Lewandowski, K., Rossaint, R. & Falke, K. J. *Int. Care Med.* **19**, 443–449 (1993).
- Puybasset, L. et al. *Int. Care Med.* **20**, 319–327 (1994).
- Rossaint, R. et al. *New Engl. J. Med.* **328**, 399–405 (1993).
- Jia, L., Bonaventura, C., Bonaventura, J. & Stamler, J. S. *Nature* **380**, 221–226 (1996).

versus rat immunoglobulin, and chromogenic localization with 3,3'-diaminobenzidine/H₂O₂. We used two methods to control specified carbonyl-dependent from carbonyl-independent (nonspecific) binding: (1) incubation of sections with 10 mM sodium borohydride (NaBH₄) for 30 min at room temperature before incubation with DNP; and (2) absorption of the antibody to DNP with 2,4-DNP-pyruvate before use.

Sections of hippocampal tissue from Alzheimer's samples and aged-matched controls show striking differences. In tissue from Alzheimer's cases, but not in controls, carbonyls were increased in neuronal cytoplasm and nuclei of neurons and glia (arrowheads in figure) indicative of a generalized increase in oxidative stress in Alzheimer's. Moreover, neurofibrillary tangles, a key pathological feature of the disease, were strongly labelled (arrows in figure), suggesting that these lesions are targeted in oxidative damage. The specificity of this technique for free carbonyls was attested by the lack of reactivity both in sections reduced by NaBH₄ and where the DNP antibody was blocked by adsorption with DNP-pyruvate (results not shown). Our results complement a previous study showing an increase in DNP-reactive material in tissue homogenates from normally aged and Alzheimer's patients⁴. The precise cellular or subcellular location of oxidative damage was not determined in this earlier study.

DNP studies using tissue homogenates will, in addition to the macromolecule-based carbonyls, also detect aldehydes arising from lipid peroxidation and glycoxidation that are displaced into the solvent by DNP-derivatization. Although the two last modifications have been described in association with neurofibrillary tangles³, the *in situ* technique described here will detect only macromolecule-based carbonyls. Our technique will thus reveal mainly carbonyls arising from side-chain oxidations, especially of lysine, arginine, proline and threonine¹, as well as carbonyls contained in lipid- or sugar-derived moieties attached to the protein by non-DNP-displacable linkages. It is probably not a coincidence that tau and neurofilament proteins, the main protein components of neurofibrillary tangles, contain an unusually high abundance of proline and lysine residues. Finally, it is worth noting that protein-bound carbonyl groups may be involved in intra- and intermolecular covalent crosslinking, the latter probably

contributing to insolubility of the tangles³. Significantly, we find disease-specific oxidative damage to nuclei, which is consistent with the increased DNA strand breakage reported in Alzheimer's disease⁵. Strand cleavage is indicative of apoptosis, but is also consistent with oxidative DNA damage.

In conclusion, the *in situ* detection of carbonyls described here is an excellent measure of oxidative stress and allows accurate determination of the extent and subcellular location of initial damage. In Alzheimer's disease, carbonyl-reactivity is not only specifically increased in association with the neurofibrillary alterations characteristic of the disease, but also is increased in neuronal cytoplasm and glial and neuronal nuclei. This is further evidence for the notion that oxidative stress, as manifested in oxidative modifications, is associated with the pathogenesis of Alzheimer disease.

Because the role of oxidative stress in the aetiology or pathogenesis of various degenerative diseases is becoming more widely recognized, this technique, in conjunction with other measures of damage, should be widely applicable.

M. A. Smith, G. Perry*, P. L. Richey
Institute of Pathology,

L. M. Sayre

Department of Chemistry,

V. E. Anderson

*Department of Biochemistry,
Case Western Reserve University,
Cleveland, Ohio 44106, USA*

M. F. Beal

*Department of Neurology,
Harvard Medical School,
Boston, Massachusetts 02114, USA*

N. Kowall

*Department of Veterans Affairs,
Bedford, Massachusetts 01730, USA*

*Senior author, to whom correspondence should be addressed.

Submarine groundwater discharge

SIR—Moore¹ infers from measurements of Radon-226 in coastal waters that submarine groundwater discharge into the South Atlantic Bight might amount to a flux equivalent to 40% of the total flow entering the sea from adjacent rivers. Noting that previous estimates of submarine groundwater discharge elsewhere in the world have generally yielded figures in the range 0.01 to 10%, Church in News and Views² discusses the wide-ranging consequences of the new results for management of water resources in coastal areas. But do the adjoining coastal aquifers receive sufficient replenishment from rainfall ('recharge') to sustain the inferred rate of submarine groundwater discharge? And is the inferred width of the zone of submarine groundwater discharge compatible with the physics of the groundwater flow system?

Exactly what is meant by submarine groundwater discharge (SGWD)? Church² defines it as "direct flow of water into the sea through porous rocks and sediments". By this definition, hydrologists recognize SGWD as a world-wide phenomenon³, the magnitude of which is dictated by local groundwater conditions. For instance, where highly permeable strata (such as cavernous limestone) underlie the coastal region, local surface runoff may be negligible and SGWD can greatly exceed the discharge from adjacent rivers.

However, when analysing the coastal hydrology along the edge of a large continent (as in the case of the South Atlantic Bight¹), then SGWD will generally equate to only a few per cent of the total discharge from adjacent major rivers, since these will typically drain catchments several orders of magnitude larger than the

recharge areas of the coastal aquifers. For this reason, the figure of 40% total SGWD deduced by Moore¹ is indeed surprising.

Moore and Church both use the term SGWD to cover *all* water flowing from the sea-bed into open water, irrespective of whether this is flow coming "from the land to the sea", or local recycling of marine water in response to tidal pumping. To eliminate confusion, I shall refer to this second usage as "total" SGWD. The distinction between the two definitions has important consequences for the wider significance of the estimated fluxes.

The rate of rainwater infiltration (recharge) to the aquifers of the coastal plain adjoining the South Atlantic Bight varies between 180 and 250 mm per yr (ref. 4). Given that the coastal plain has a reasonably uniform breadth of around 180 km, then the total annual recharge in the region is equivalent to average fluxes in the range $(2.8-4) \times 10^{10}$ litres per day (1 d^{-1}). Over the long term, the same quantity of ground water must discharge, both to inland rivers and to the sea, as SGWD. Mathematical modelling exercises⁴ suggest that the proportion of groundwater discharge going to inland rivers amounts to as much as 97% of the total recharge; consequently, the amount of water available for SGWD probably does not exceed $1.2 \times 10^9 \text{ l d}^{-1}$. This is only 4% of the total SGWD flux of $3 \times 10^{10} \text{ l d}^{-1}$ inferred by Moore¹, and equivalent to only 1.7% of the total fresh water discharged to the South Atlantic Bight from rivers during the study period.

Moore deduces that the ²²⁶Ra-enriched waters he measured were discharged from a fresh saline groundwater mixing zone, where sea water invading sediments that

1. Levine, R. L., Williams, J. A., Stadtman, E. R. & Shacter, E. *Meth. Enzym.* **233**, 346-357 (1994).
2. Stadtman, E. R. *Rev. Biochem.* **62**, 797-821 (1993).
3. Smith, M. A., Sayre, L. M., Monnier, V. M. & Perry, G. *Trends Neurosci.* **18**, 172-176 (1995).
4. Smith, C. D. *et al. Proc. natn. Acad. Sci. U.S.A.* **88**, 1043-1054 (1991).
5. Su, J. H., Anderson, A. J., Cummings, B. J. & Cotman, C. W. *Neuroreport* **5**, 2529-2533 (1994).

were previously occupied by fresh ground water causes desorption of divalent metals. Thus, the total SGWD flux estimated by Moore² necessarily includes a substantial marine component. From the estimate of SGWD given above, however, the most dilute mixture emanating from the mixing zone would still comprise more than 90% sea water. This implies either that the total SGWD flux estimated by Moore is too high, or that the bulk of the total SGWD is indeed of marine origin. If the latter is the case, then Moore's deduction that SGWD is the equivalent of 40% of river flow arises from a failure to compare like with like. In other words, the SGWD entering the South Atlantic Bight is unremarkable, and would not in itself prompt a reappraisal of the oceanic chemical mass balance, or of water resources engineering practice.

Turning to my second point, that of the width of the zone of SGWD: using inverse deduction from the spatial distribution of ²²⁶Ra in the near-shore waters of the South Atlantic Bight, Moore infers that the zone of total SGWD extends up to 20 km out to sea. How does this estimate compare with independent evaluations based on groundwater flow around fresh-saline water interfaces?

Density-dependent flow has been analysed for more than a century using simplified analytical solutions^{3,5}, and more realistic solutions are increasingly being obtained by numerical methods⁶. Such studies demonstrate that the distance (D_{SGWD}) out to sea over which SGWD generally persists can be expressed (for the case of a homogeneous aquifer at steady-state)³ as: $D_{SGWD} = (20q') / K$, where q' is SGWD per unit width of coastline and K is aquifer hydraulic conductivity (permeability with respect to water).

Scoping calculations show that, for the overwhelming majority of cases, D_{SGWD} will never exceed a few hundred metres. In the case of the South Atlantic Bight, we can estimate the likely maximum value of D_{SGWD} by combining a minimum estimate of local hydraulic conductivity⁴ (10 m per day) with an upper-bound q' value (93,750 l d⁻¹ m⁻¹, obtained by dividing Moore's value of 3×10^{10} l d⁻¹ by 320 km). This yields a maximum estimate for D_{SGWD} of 188 m. Although this agrees well with direct observations of D_{SGWD} elsewhere on the eastern seaboard of the United States⁷, it is two orders of magnitude smaller than the 20 km suggested by Moore.

Whatever the width of the zone of SGWD, it is necessary also to consider the likely width of the zone of fresh-saline water mixing which separates the terrestrial and marine ground waters. Direct observations of such zones reveal that they are typically a few metres wide, and seldom exceed a few hundred metres⁷.

Consequently, D_{SGWD} probably underestimates the width of the zone of total SGWD by a very small margin. In other words, it is physically unreasonable to postulate a mixing zone of 19,812 m extending seawards from a 188-m zone of SGWD to make a total SGWD discharge belt 20 km wide, as proposed by Moore.

Paul L. Younger

*Department of Civil Engineering,
University of Newcastle,
Newcastle Upon Tyne NE1 7RU, UK*

1. Moore, W. S. *Nature* **380**, 612–614 (1996).
2. Church, T. M. *Nature* **380**, 579–580 (1996).
3. Fetter, C. W. *Applied Hydrogeology* (Macmillan, London, 1994).
4. Back, W., Rosenshein, J. S. & Seaber, P. R. (eds) *Hydrogeology, Geology of North America Volume O-2* (Geol. Soc. Am., Boulder, 1988).
5. Ghybens W. B. *Tijdschrift van Het Koninklijk* (Inst. van Ing. Amsterdam, 1888).
6. Aliwi, A. S. thesis, Univ. Newcastle (1993).
7. Kohout, F. A. *J. geophys. Res.* **65**, 2133–2141 (1960).

MOORE AND CHURCH REPLY — Younger questions whether there is adequate terrestrial groundwater recharge to sustain significant fluxes of fresh ground water to the ocean. He argues that almost all of the recharge is captured by streams and little remains to discharge to the ocean. The reference he cites¹ for infiltration of 180–250 mm per yr states that these estimates are derived from an analysis of the base flow component of stream flow. In other words, if almost all the infiltration is captured by streams, 180–250 mm per yr are required to maintain stream flow at observed levels. Because the streams are assumed to capture almost all the infiltration, this assumption drives the infiltration models in a circular argument.

We question this assumption and maintain that there is ample uncertainty in estimates of water balance and associated below-ground inventories to allow significant fresh groundwater fluxes to enter the ocean. Younger should also be aware that the catchment area of the major rivers draining into the South Atlantic Bight is controlled by the Appalachian divide and is only about a factor of three larger than the catchment area of the coastal aquifers, not several orders of magnitude as he estimates.

The answer to this question also depends on your viewpoint. To a terrestrial hydrologist, an annual loss of a few per cent of fresh ground water to the ocean may seem trivial. But, to a chemical oceanographer, discharge of this amount of fresh water with larger volumes of salty submarine ground water (SGWD) may cause a re-evaluation of chemical inputs to the ocean². Therefore, we must take issue with Younger's opinions that even if SGWD is largely of marine origin, its input to the ocean is "unremarkable". The SGWD has chemical characteristics remarkably different from fresh groundwater or ocean water³. The material flux to the ocean derived from SGWD is the product of its volume times the difference

in concentration from ocean water.

If SGWD to the South Carolina coast is of the order of 3×10^{10} l d⁻¹ (ref. 3), it must contain about 100 times more ²²⁶Ra than surface ocean water; if the flow is only 10% of that estimated, the ²²⁶Ra enrichment must be a factor of 1,000, much higher than we have measured⁴. Flow of large volumes of this chemically altered water may have an enormous impact on coastal ocean chemistry and productivity². As we and others have shown^{3,6}, SGWD also carries high concentrations of Ba, P, N, Mn, Fe, Cu, Zn, Cd, Pb and Ni. The power of the Ra tracer in quantifying the regional effect of SGWD lies in the conservative behaviour of ²²⁶Ra once it is injected into the coastal ocean. Thus, ²²⁶Ra fluxes offshore should serve as a proxy for inputs of other elements via SGWD.

Younger's second point concerns the width of the zone of SGWD. The input width of 20 km was chosen to correspond to a strong pycnocline (density gradient in the water column) during the period of investigation³. Input of ²²⁶Ra below the pycnocline from SGWD further off shore was not considered in the mass balance. Simmons⁴ measured substantial fluxes ($2\text{--}20$ l m⁻² d⁻¹) of SGWD at sites 15–37 km from the South Carolina coast. He also measured fluxes of ocean water into the sediments in this area. We have recently found large ²²⁶Ra enrichments at a site 60 km from shore. Drilling has confirmed the presence of brackish water beneath the continental shelf all along the east coast of the United States^{7,8}. Clearly, SGWD is occurring throughout the continental shelf, not just from the inner 0.2 km as Younger calculates. When we quantify SGWD from these deeper sites, its impact on the coastal ocean will have to be revised upward.

The challenge to hydrologists is to develop SGWD models that include tidal pumping due to diurnal, monthly, seasonal or longer changes of sea level; saltwater intrusion and changes in groundwater usage; mixing and chemical reactions within coastal aquifers; and the expulsion of a saline end-member.

Willard S. Moore

*Department of Geological Sciences
University of South Carolina
Columbia, South Carolina 29208, USA*

Thomas M. Church

*College of Marine Studies
University of Delaware
Newark, Delaware 19716, USA*

1. Back, W., Rosenshein, J. S. & Seaber, P. R. (eds) *Hydrogeology, Geology of North America Volume O-2*. (Geol. Soc. Am., Boulder, 1988).
2. Church, T. M. *Nature* **380**, 579–580 (1996).
3. Moore, W. S. *Nature* **380**, 612–614 (1996).
4. Simmons, G. M. *Mar. Ecol. Prog. Ser.* **84**, 173–184 (1992).
5. Shaw, T. J. *et al.* *EOS* **76**, F318 (1995).
6. Krest J. M. & Moore W. S. *EOS* **77**, 5153 (1996).
7. Hathaway, J. C. *et al.* *Science* **206**, 515–527 (1979).
8. Manheim, F. T. *Trans. N. Y. Acad. Sci.* **2**, 839–853 (1967).

Corrupted by money?

Michael V. Posner

The Economic Laws of Scientific Research. By Terence Kealey. Macmillan/St Martin's: 1996. Pp. 382. £45, \$75 (hbk); £15.99 (pbk).

TERENCE Kealey is a clinical biochemist of good reputation at the University of Cambridge, England. In his first chapter he establishes his continuing theme, suggesting that poor Francis Bacon set us on the wrong lines: "Where Bacon believed that [the origins of technology] would flow from academic science, Adam Smith maintained that it largely derived from the industrial development of pre-existing technology". Kealey then gives us half a dozen chapters of potted economic history, much of it amusing and provocative, illustrating the folly of the Baconian model and the correctness of Smith — hence, no doubt, Kealey's appeal to "economic laws". I neither accept nor deny his interpretation of the past (although I find it hard to believe that Smith was as dogmatic as Kealey suggests); but this is all a prelude to Kealey's manifesto for the present and the future, which makes up the rest of his book.

In chapter 10 he writes from his own experience, and illuminates sceptically some accepted UK doctrines about pharmacological innovation. Elsewhere (chapter 9 — "Why the Linear Model fails") he deploys familiar arguments to throw cold water on those simple cause-and-effect stories which suggest that more extensive subsidy of research would yield more useful technology: Smith right, Bacon wrong, is the theme. And he uses statistics from the Organization for Economic Cooperation and Development (in chapter 10, and throughout the second half of the book) to suggest that it is very easy to waste research funds, especially when financed by taxpayers' money, and that it is easier still to spend far too much on research. (The former Soviet Union is a good example, paradoxical though this might seem to the many top-class Western scientists who found Soviet research work in their fields fascinating and valuable.)

All this is good fun, and the more it stings, the more good it does: it makes us all stop and think. He could profitably go even further, and argue that research is not 'an input' into economic activity (creating further prosperity for the citizens), but 'an output' (like Scotch whisky or visits to the grand opera) that is pleasurable to indulge in but which absorbs economic resources rather than adding to them. Scientists, he could suggest, not only want to make excessively expensive journeys on

the backs of new technologies (nuclear energy, space exploration — there are good passages on the US experience in these fields in chapter 8), but also, like coal-miners who lose their jobs in the face of competition from new technologies (such as gas turbines, which arrived unexpectedly on a fast track of industrial development unaided by subsidy), demand



Folly or foresight? — Francis Bacon (left) believed that new technology would flow from academic science, whereas Adam Smith (right) argued that it arose largely from the industrial development of pre-existing technology.

the maintenance of traditional lines of research because they are 'basic'.

The trouble is that Kealey is not content to act as a gadfly; he really aims to influence the course of events directly and wants everything, the world over, to be organized differently. The appeal (not merely in his title, but throughout) to 'economic laws' is a rather chilling part of his campaign. There are indeed some economic principles, and some do offer a little support to the author: if you subsidize (from taxes) an activity, you are likely to get more of it; if you subsidize an activity on which other people are spending their own money, you risk using tax-based subsidies to 'crowd out' private expenditure; if you push large long-term projects in your own country (as in the United States, the United Kingdom or France), and finance them by taxes, other countries will be happy to take the free

ride, but will not chip in to pay for the failures. But these are best regarded not as 'laws' but as wise warnings from experienced observers. Kealey seems to think that, like Newton, he has discovered great new principles that instruct him indubitably that the world must be turned upside down. For instance, he asserts in a postscript that there is an economic law, which he reinforces by short quotations from Hume and Solzhenitsyn, that requires us to abandon publicly funded education. The true position is that there are indeed economic arguments that point to the benefits from such policies, but also economic arguments that point the other way, and facts that can be deployed on both sides: appeal to 'laws'

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Mary Evans

the promotion of technology *may* be useful, but it is preeminently the price we pay to obtain also some free, basic money: it is the sop to Cerberus, the Danegeld, or (with increased desperation) the meat you throw to the wolves who are chasing your sledge across the tundra.

But why do governments insist on pushing university scientists into 'useful' channels? Here Kealey deploys all his anti-Baconian arguments. Successful projects will make money for the innovator (and his close followers). Isn't it best to leave with these innovators, in the private sector, the task of choosing the best (most promising) projects, and to control and finance them? The author examines quite closely and carefully some of the fundamental papers that argue the contrary case, and it is not easy to demolish his conclusions (such as his analysis of Arrow's case in chapter 9). He has good fun in attacking the big public spenders in Europe and the United States. His appeal to the Japanese experience is insufficiently substantial, but carries some weight. So there needs to be some answer to Kealey on this point. Many government papers, in many countries, attempt to formulate such an answer: Kealey does them scant justice — indeed, he largely ignores them, no doubt because they start from a standpoint he does not find congenial.

To establish rational discussion with Kealey, we could start, open-mindedly, by acknowledging that the distribution of researchers between academic institutions and industry may be wrong — perhaps it would have been better to run all universities like high schools, staffed by full-time teachers: but, as a matter of fact, in many (not all) fields of science, the majority and the best of scientists are in universities or nonprofit institutes. We could starve them out, and push them onto the labour market (as is happening in the former Soviet Union). Kealey flirts with this approach, although, appropriately for a Cambridge don, he is a little coy in recommending the end of the university as he knows it. If we do not choose such creative destruction, public money has to grease the wheels, provide the interface, keep the research machine ticking over. Such public money serves the function of allowing the Kealey mechanism (what he would call the 'economic laws') to operate — essentially, there should develop a new dual-support mechanism, with public money supporting an active and vigorous research base, and private money spending to use that base, employ its graduates and help establish its agenda.

It is to Kealey's credit that he enlivens this debate. It is perhaps natural that he does so in a boisterous Punch and Judy style. It is undesirable that he sometimes also sets up Aunt Sallies to belabour (for example, in his over-extended polemic against Martin and Irvine in chapter 11).

It is irritating to this reviewer that he scatters names and references around as if he were determined to retrieve every entry from his database. But large bits of the book are a good read. □

Michael V. Posner, former secretary-general of the European Science Foundation, is at Rushwood, Jack Straw's Lane, Oxford OX3 0DN, UK.

New shapes from old genes

Jonathan Slack

The Shape of Life: Genes, Development, and the Evolution of Animal Form. By Rudolf A. Raff. *University of Chicago Press: 1996. Pp. 520. \$55, £43.95 (hbk); \$29.95, £23.95 (pbk).*

In modern biological science, development and evolution represent almost totally separate 'continents of knowledge'. Those who investigate development are interested in the mechanism of how things work. They use a small repertoire of model organisms, each of which has its own peculiar technical advantages, and they are concerned above all with isolating and understanding the genes that control development.

Many evolutionary biologists tend to be relatively uninterested in actual molecular mechanisms. They want to know things such as whether sympatric speciation is possible, how altruism could evolve or how sexual reproduction can be maintained despite the apparent cost in resources. They just assume that in some way genes control phenotype and in their speculations feel free to postulate a gene for anything without worrying too much about the known molecular reality.

On the face of it, this great epistemological gulf seems unhelpful. Developmental biologists are, after all, uncovering the mechanisms by which genes are translated into anatomy. They should be in a good position to tell us exactly what has happened in evolution in terms of the genetic changes that underlay a particular anatomical change. This has the potential to answer some of those big questions of macroevolution, such as "why did all the animal body plans appear in a short period in the early Cambrian?" or "why are there no six-legged vertebrates?" or "what is the real difference between a man and a fish?"

Over the years there have been sporadic efforts by a few heroic individuals to unite the two fields, and *The Shape of Life* is an important work in this tradition. The author, Rudi Raff, is best known for his work on sea-urchin development and how it has become modified in the subset of sea-urchin species that develop directly to

the adult, leaving out the stage of the pluteus larva. The book covers all the areas in which development and evolution meet: animal phylogeny, the Cambrian explosion, cladistics, molecular phylogeny, the phylo-typic stage, heterochrony, developmental constraints on variation and the genetic control of developmental modules. It is written in a very informal style, with references given implicitly rather than formally and much use of the first person and its contractions 'I'll' and 'I've'. This may be irritating to some, but it does have the overall effect of making a long book easier to read than it might have been. Key themes are usually introduced by some anecdote, such as the long sledge journey in 1911 to recover eggs of the emperor penguin from a remote spot in Antarctica, or human-animal mosaics depicted in Cro-Magnon cave paintings. These can be entertaining and break up the text in a useful way.

I can mention here only a tiny handful of the many interesting questions discussed. For example, the explosion of new phyla in the early Cambrian has no parallel even following the mass extinctions at the end of the Permian and Cretaceous periods. In the Permian extinction, it is estimated that 95 per cent of marine species disappeared, and yet no new phyla evolved to fill all the vacant niches. This suggests that even the few survivors covered enough 'morphospace' to take advantage of the available opportunities. Among the modern insects, where the molecular basis of development is now best understood, a comparison of the fruitfly *Drosophila* with four-winged types has shown that wings on the third thoracic segment are not produced by an absence of expression of the *Ubx* gene but rather by a modification of the responses to codes in *Hox* gene clusters. This example should warn developmental biologists against thinking that because a certain transformation can be produced experimentally, it must have happened that way in evolution. Another randomly selected point that grasped my attention was the significance of the primitive fossil amphibian *Acanthostega*, which has eight digits on its forelimb. We have always considered five digits to represent the primitive condition and no cladistic analysis of living vertebrates could possibly predict anything different.

I have a few minor gripes: Raff occasionally confuses homeobox and *Hox* genes; in chapter 4 he manages to discuss the homology of annelid and arthropod segmentation mechanisms without mentioning the *engrailed* gene; and in chapter 10 he mentions two *Hox* gene clusters in cephalochordates (there is only one). But in general the book can be recommended to students and perhaps even the biologically educated general public as a goldmine of fascinating information about macroevolution.

In 1983, the author wrote a book with T. Kaufman entitled *Embryos, Genes and Evolution*, which covered some of the same general problems. Raff must have felt that we now know so much more about development that a completely new book was necessary. In the intervening years, developmental biologists have certainly discovered some intriguing primitive features of animal body plans such as the common genes controlling antero-posterior and dorso-ventral axis formation, the eye and heart territories. Three years ago, some of us ventured to call this common constellation of gene expression domains 'the zootype' (see *Nature* 361, 490; 1993), and the concept does seem to have been strengthened with each new discovery. These results do at least tell us that animals really are monophyletic (that is, they are descended from a single ancestral species), something that has not always been held as self-evident by phylogenetic theorists.

On the other hand, when we come to variation and the origin of derived features, developmental biology still seems unequal to the task before us. We still do not really know what sort of constraints on viable mutational variation are imposed by the developmental programme. We still do not understand the nature of the changes in the programme that have caused particular evolutionary morphological transformations, such as those from indirect- to direct-developing sea urchins. We still have no idea whether an elephant fly could, as Raff provocatively asks, evolve into an elephant. From this point of view the book, despite its many positive features, may still be somewhat ahead of its time. □

Jonathan Slack is in the Developmental Biology Programme, School of Biology and Biochemistry, University of Bath, Bath BA2 7AY, UK.

For whom the net tolls

Steve Fuller

Scholarly Publishing: The Electronic Frontier. Edited by Robin Peek and Gregory Newby. MIT Press: 1996. Pp. 363. \$35, £29.50.

INVETERATE 'netsurfers' typically believe that electronic communication has removed all the material obstacles that have traditionally prevented scholarship from enjoying universal access and immediate impact. If you are one of those people, the odds are that your computer is connected to a university or corporate mainframe, which means that you are not directly charged any users' fees and you rarely suffer from delays in transmission.



THE major aurora of 24–25 March 1991, visible from the south of England. From the cover of the second edition of *The Aurora: Sun–Earth Interactions* by Neil Bone (Wiley, £19.99 (pbk)), which contains a colour mini-atlas of auroral types and forms, descriptions of the different types of aurora, information on related high-atmosphere phenomena and practical advice on auroral observation. For a review of the first edition, see *Nature* 360, 428 (1991).

However, if you log on through a modem connected to your telephone, the 'information superhighway' fast loses its reputation for being a frictionless medium of thought. Depending on your location, telephone access fees can mount up very quickly, and the rate structure of commercial electronic carriers may force you into a suboptimal service that frequently suspends your messages in limbo while those of premium customers zip back and forth. Finally, if you are a technophobic scholar lacking connections to a major institution, the Internet merely widens the gap between you and the rest of the intellectual world.

These homely observations provide an anchor in reality for this book's wide-ranging proposals to embed intellectual life in the electronic medium. About half of the 19 articles project electronic utopias tailor-made for a preferred constituency, usually academics, librarians and publishers. The other half offer more sober and useful analyses of the medium by lawyers, economists and information scientists. Perhaps the best way to read this book is to use the essay by Rob Kling and Roberta Lamb as a guide to understanding the other essays. Kling and Lamb systematically contrast the rhetorics invoked by 'utopian' and 'anti-utopian' visionaries of the new electronic age. Kling has recently edited a lively special issue of *The Information Society* (vol.

11, no. 4, 1995) that manages to get some of the competing visionaries to look each other straight in the eye.

Not surprisingly, the utopians have so far received the most media coverage in their attempt to shape the future of scholarship on the Internet. One imagines them also to be among the inveterate netsurfers. Stevan Harnad, Andrew Olydzko and Ann Okerson clearly fit into this mould; but so too does Jean-Claude Guédon, a literary historian who celebrates the new intensity and cross-fertilization that the Internet brings to scholarly communication. As they see it, scholars have a clear sense of what they want and sufficient resources to make it happen, even though a full transition to the electronic medium may incur some short-term institutional costs. The utopians envisage business-as-usual in academic institutions becoming increasingly efficient as scholars are able to access materials and audiences more easily, produce more text and receive quicker feedback. They never consider the congestion problems that will arise as more scholars log on, nor do they worry much about assigning intellectual property rights to tracts in cyberspace. In fact, there is a tendency among the utopians to blame the publishing industry for anything that reminds academics that their work has an economic dimension.

No such reminder is needed for those

PRISCACARA oxyprion, a common perch-like fish that lived around 50 million years ago in lakes of what is now central North America. Several hundred fossil species of perch-like fishes have been discovered, and some of the best can be seen in the gallery on vertebrate origins that opened last month at the American Museum of Natural History in New York. Its opening marks the completion of the museum's seven-year redesign, restoration and replenishment of its fossil halls. Of the 600 specimens on display, around 85 per cent are genuine fossils, not casts — an unusually large proportion for any collection.



who were originally captivated by the utopian promises and now face mounting access charges as they 'keep up with the Joneses' in cyberspace. In one of the most informed pieces in the volume, Ira Fuchs, the founder of BITNET, observes that many of the proposed pricing schemes for the Internet unwittingly hit financially vulnerable institutions the hardest, especially primary and secondary schools. This is because an unusually large percentage of the transmissions made by schools contain multimedia imagery, which travel in large packets that can easily congest the Internet, regardless of the speed at which they are transmitted. Encouraged by both public and private sector investment in the 1980s and early 1990s, schools built substantial information infrastructures that ultimately aimed to meet all their instructional needs. At that time, school administrators were led to believe that government would either continue subsidizing the information infrastructure or would regulate the markets in which it is transacted. But the advent of Internet privatization has put this tacit commitment in doubt, leading schools to think twice about their own commitment to the electronic medium.

The role of government in our electronic future is the most sensitive issue faced by the volume's contributors. Will the Clipper Chip turn government into 'Big Brother', or will an automated civil service permit ordinary citizens to dispense with government bureaucrats altogether? Realistically, however, apart from the role of utility regulator that Fuchs suggests, governments will probably be forced to extend the scope of intellectual property law into arenas traditionally reserved for peer-reviewed processes. Brian Kahin, an attorney affiliated with Harvard's Kennedy School of Government, argues that the relative ease of electronic communication often makes it difficult to decide who deserves credit for

an idea once it has been embedded in hypertext and circulated through cyberspace. Nevertheless, the lack of clear credit assignments threatens to undermine the norms of intellectual life as scholars move increasingly on-line. Our utopian visionaries may object that intellectual life would be better off were it not parcelled out into property rights, but we have yet to see what the alternative would look like in practice. □

Steve Fuller is in the Department of Sociology and Social Policy, University of Durham, Durham DH1 3JT, UK.

Mature offerings

Michael Proctor

The Anther: Form, Function and Phylogeny. Edited by William G. D'Arcy and Richard C. Keating. Cambridge University Press: 1996. Pp. 351. £55, \$80.

POLLINATION biologists, palaeobotanists and developmental morphologists may view anthers very differently — and in any field it is all too easy not to look outside one's own speciality. It is perhaps not so much that, in the words of the editors, "the anther [has] received relatively little scientific attention" as that the attention has been divided and there has been too little communication across the divisions. The number of references cited in this book seem to confirm this.

The Anther grew from a symposium at the 1993 International Botanical Congress at Yokohama, but is much more than a routine conference volume. D'Arcy's opening chapter provides an excellent and wide-ranging introduction to the structure and function of stamens. The book concludes with a review by Keating of methods for anther investigations (from optical and

electron microscopy, histochemistry and DNA techniques, to palynology and pollination biology), and a bibliography compiled by Lynch and Gregory of nearly 1,500 references on stamen morphology, indexed by subject and plant family. In between are chapters on the evolution of stamens, which draw on both fossil evidence (Crepet and Nixon) and cladistics (Hufford); diversity in angiosperm stamens (Endress); anther adaption in relation to pollination (Bernhardt); and various other topics.

All in all, it is an immensely stimulating collection. Crepet and Nixon's account of recently studied fossil flowers from Turonian deposits in North America enlarges our ideas about Cretaceous angiosperm radiation; the small size of many of the flowers they describe will surprise plenty of readers. Cladistics is useful in making systematic inferences from comparative data, but the conceptual tree structure inherent in the method has its hazards when at any period in geological time we must be looking at much parallel development in diverse floras. Are stamens and carpels homologous (that is, of common descent)? Burger argues persuasively that they are not, and that laminar stamens in Magnoliales are secondary characters. And what is the calcium-oxalate package doing in Solanaceous (and a few Ericaceous) anthers?

The Anther is a splendid source of both information and ideas, well edited and nicely produced. It touches on several areas of angiosperm biology where we need to think laterally, and perhaps to abandon some comfortable but outworn intellectual ideas and recognize that the fossil history of the angiosperms and pollination biology (and the role of the anther in both) are more complex, more interesting and more challenging than we have tended to assume. □

Michael Proctor is in the Department of Biological Sciences, University of Exeter, Exeter EX4 4QG, UK.

Late Proterozoic rise in atmospheric oxygen concentration inferred from phylogenetic and sulphur-isotope studies

Donald E. Canfield & Andreas Teske*

Max Planck Institute for Marine Microbiology, Celsiusstrasse 1, Bremen 28359, Germany

The evolution of non-photosynthetic sulphide-oxidizing bacteria was contemporaneous with a large shift in the isotopic composition of biogenic sedimentary sulphides between 0.64 and 1.05 billion years ago. Both events were probably driven by a rise in atmospheric oxygen concentrations to greater than 5–18% of present levels—a change that may also have triggered the evolution of animals.

ANIMAL life (metazoans) evolved late in Earth history. Amino-acid sequence comparisons indicate an evolutionary radiation 0.8 to 1.0 Gyr ago (billion years before present)¹, although the Ediacaran fauna at 0.59 Gyr (ref. 2) represent the earliest well-accepted fossil evidence. The late emergence, and subsequent rapid radiation, of metazoans through the terminal Proterozoic era (0.544 Gyr (ref. 3)) is thought to have coincided with an increase in atmospheric oxygen concentration from less than 3–10% to up to 100% of present atmospheric levels (PAL)^{4–6}. This increase is believed to have provided sufficient oxygen for animal respiration and for the synthesis of the structural protein collagen, which is universally distributed among metazoans⁷.

Oxygenation of the Earth's surface occurs as organic carbon is buried in sediments, releasing oxidized equivalents ultimately as ferric iron, sulphate and atmospheric oxygen^{8,9}. Stable carbon isotope studies indicate high organic-carbon burial rates between 0.9 and 0.6 Gyr (refs 8,10), supporting the proposed rise in atmospheric oxygen in the late Proterozoic. No additional evidence, however, has been available to verify this rise in oxygen. Here we report a significant evolutionary radiation of non-photosynthetic marine sulphide-oxidizing bacteria, and a dramatic increase in sulphur-isotope fractionation at between 0.64 and 1.05 Gyr. We argue that both events were driven by an increase in atmospheric oxygen from levels less than 5% to 18% PAL to greater concentrations.

Bacterial evolution

The phylogeny of the prokaryotes, as indicated by 16S ribosomal RNA sequences, demonstrates major evolutionary radiations of bacteria with similar metabolisms or environmental niches¹¹. In general, major bacterial groups are thought to originate through rapid, episodic evolution, in response to environmental change¹¹. The oxygenation of bacterial habitats is such a change, and triggered the relatively late emergence of aerobic and microaerophilic bacteria throughout the eubacterial kingdom^{11,12}. We focus here on the radiation of non-photosynthetic sulphide-oxidizing bacteria of the γ -proteobacterial subdivision, emerging near the basal transition of the γ and the β subdivisions¹¹. This radiation produced the largest cluster of aerobic and microaerophilic sulphide oxidizers among the Eubacteria; we suggest that the timing of this radiation constrains the oxygenation of sulphide-containing bacterial, particularly marine, habitats. Sulphide oxidizers are of interest here because the occurrence of obligate

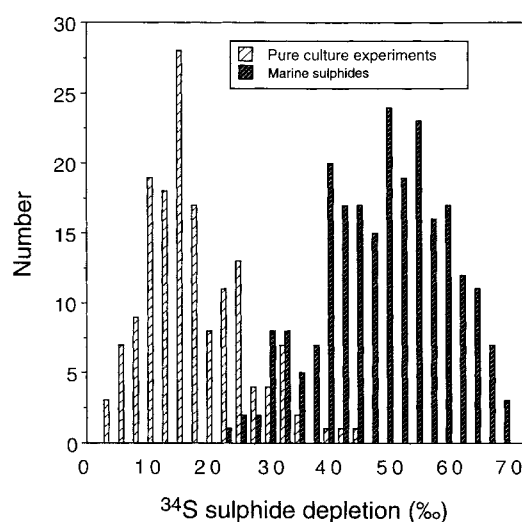


FIG. 1 Histogram of ³⁴S sulphide depletions for marine sedimentary sulphides and for sulphides generated by sulphate reducing bacteria. For sediments, ³⁴S depletions are taken as the difference between the isotopic composition of pore-water sulphate and sedimentary sulphide. Isotope values are generally from near the sediment–water interface where the isotopic composition of sulphate is near to the seawater value. For each sediment only the maximum ³⁴S depletion measured is reported. Maximum values should be mostly free of the ‘closed system’ addition of ³⁴S-enriched sulphide, and hence best represent fractionations during the biological processing of sulphur. Experimental ³⁴S depletions are the difference between the isotopic composition of sulphate and the sulphide produced by pure cultures of sulphate-reducing bacteria. All isotopic compositions were originally presented as ‰ (parts per thousand) differences relative to Canyon Diablo troilite (CDT), such that

$$\delta^{34}\text{S}_{\text{sam}} = \left\{ \left[\left(\frac{{}^{34}\text{S}}{{}^{32}\text{S}} \right)_{\text{sam}} - \left(\frac{{}^{34}\text{S}}{{}^{32}\text{S}} \right)_{\text{CDT}} \right] / \left(\frac{{}^{34}\text{S}}{{}^{32}\text{S}} \right)_{\text{CDT}} \right\} \times 10^3$$

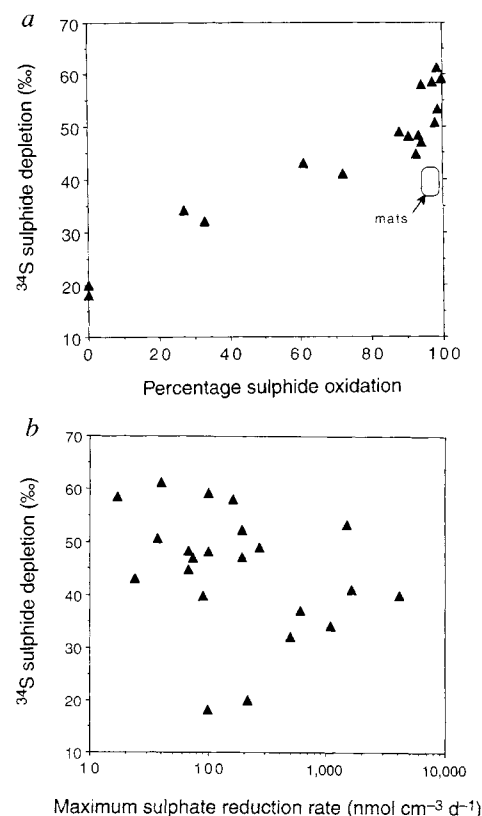
where sam is the sample, and $\delta^{34}\text{S}_{\text{sam}}$ is the original isotopic composition of the sample. A complete list of references and data are available as Supplementary Information.

sulphide-oxidizing bivalve symbionts, within the same proteobacterial subdivision, allows us to estimate the timing of the radiation event.

We consider the microaerophilic filamentous sulphur bacteria *Beggiatoa*, *Thiothrix* and *Thioploca*, and the aerobic to microaerophilic non-filamentous *Thiomicrospira* and *Thiobacillus*, which

* Present address: Woods Hole Oceanographic Institution, Biology Department, Woods Hole, Massachusetts 02543, USA.

FIG. 2 a, ^{34}S sulphide depletion for sedimentary sulphides (see Fig. 1) compared to the percentage of sulphide produced by sulphate reduction that is lost from the sediment ($R^2 = 0.888$). Sulphate reduction rates have, in all cases, been directly measured with radiotracer. Sulphide loss is calculated as the percentage difference between the depth-integrated rate of sulphate reduction and the rate of sulphur burial (as pyrite and acid-volatile sulphide phases) in the sediment. Fractionation values at zero sulphide oxidation for sediment from the mouth of the Weser estuary are from published experimental sediment incubations⁵⁶ and this study. Data for the sediment from the Belt Seaway, northern Denmark, the Chilean upwelling zone, and Golfo Dulce, Costa Rica are from ref. 32 and this study. Also included is a region representing the isotopic composition and percentage sulphide oxidation for cyanobacterial microbial mats from Baja California, Mexico, Solar Lake, Sinai, and the photosynthetic marine sapropel at Mangrove Lake, Bermuda. In modern photosynthetic sediments, sulphide is oxidized by both phototrophic pathways (not requiring oxygen) and non-phototrophic pathways (requiring oxygen). In non-photosynthetic sediments, sulphide is oxidized either directly with oxygen, or indirectly by reaction with Fe and Mn oxides and with nitrate. These oxidants are generated, and regenerated by oxidation of their reduced forms (Fe^{2+} , Mn^{2+} and NH_4^+) with oxygen. b, ^{34}S sulphate depletion compared to maximum rate of sulphate reduction rate measured by radiotracer in marine sediments. Maximum sulphate reduction rate is taken as a comparative measure of sulphate-reducer metabolic activity in the sediment region where the diagenetic formation of sulphides is generally most significant, and for which most of the isotopic compositions in a have been obtained. An $R^2 = 0.023$ demonstrates a lack of significant correlation. Data are from the same sediments as in a, with a few additional sites for which percentage sulphide oxidation could not be determined.



all oxidize sulphide or sulphur with either oxygen or nitrate as the electron acceptor. *Thiothrix* grows in tufts attached to pebbles, rocks and plant material, in sulphide- and oxygen-containing flowing water¹³. *Beggiatoa*, a common constituent of modern microbial mats and sulphidic sediments, is typically found in steep opposing gradients of O_2 and H_2S , with the bacteria confined to the oxic-anoxic interface¹⁴. The denitrifying sulphide oxidizer *Thioploca* shuttles between spatially separated pools of sedimentary sulphide and seawater nitrate. *Thioploca* mats occur in low-oxygen waters ($< 5 \mu\text{M O}_2$) of the Peru–Chile upwelling region¹⁵, where nitrate is found at levels of $25 \mu\text{M}$. Nitrate is formed from the aerobic decomposition of marine organic matter—approximately one mole nitrate is produced for every 8.6 moles oxygen

molecules used—and thus the metabolism of *Thioploca* is indirectly coupled to an aerobic environment. *Thiomicrospira* spp. have been isolated from sulphidic marine sediments and hydrothermal vent waters and crusts¹⁶, and can grow in sulphide/oxygen gradients, but are not bound to these. *Thiobacillus* spp. are also not bound to chemical gradients, and occur commonly in both water columns and sediments, using oxygen or nitrate in sulphide oxidation¹⁶.

We used 16S rRNA to determine a time frame for the evolutionary radiation of these sulphide oxidizers. 16S rRNAs occur ubiquitously in all prokaryotes, and remain highly conserved in their primary sequence and secondary and tertiary structure, owing to functional coupling with the evolution of the ribosome.

TABLE 1 Average 16S rRNA sequence distances

	(1)	(2)	(3)	(4)	(5)	(6)
(1) Endosymbiotic bacteria Vesicommyidae–Mytilidae	—	13.11 ± 1.20	13.24 ± 0.32	14.52 ± 1.26	15.26 ± 0.92	20.33 ± 0.92
(2) Endosymbiotic bacteria Lucinidae–Thyasiridae		—	12.26 ± 0.90	13.29 ± 1.51	13.82 ± 0.94	17.88 ± 1.18
(3) <i>Thiothrix</i>			—	14.65 ± 0.94	16.54 ± 0.88	17.08 ± 0.79
(4) <i>Beggiatoa</i> – <i>Thioploca</i>				—	17.58 ± 0.64	16.68 ± 1.37
(5) <i>Thiomicrospira</i>					—	22.26 ± 0.59
(6) <i>Thiobacillus</i>						—

Average distances ($\pm$ standard deviations) (percentage substitution per site) between deep-branching lineages of γ -subdivision non-photosynthetic sulphide-oxidizing bacteria, calculated using several bacterial species for each lineage, as follows: (1) Bacterial endosymbionts of *Vesicommya chordata*, *Calyptogena magnifica*, *Calyptogena elongata*, *Calyptogena* species (Oregon) and *Bathymodiolus thermophilus*; (2) Bacterial endosymbionts of *Thyasira flexuosa*, *Anodonta philippiana*, *Codakia orbicularis*, *Lucina floridana*, *Lucinoma aequizonata* and *Lucinoma annulata*; (3) *Thiothrix nivea* and *Thiothrix ramosa*; (4) *Beggiatoa alba*, *Beggiatoa* species (Strain 1404–13 Göttingen) and *Thioploca ingrica*; (5) *Thiomicrospira pelophila*, *Thiomicrospira crunigena* and *Thiomicrospira* species (hydrothermal vent strain MA2-6); (6) *Thiobacillus caldus*, *Thiobacillus ferrooxidans* (strain N-Fe4) and *Thiobacillus thiooxidans* (strain B-S3) (*Thiobacilli* of the β - γ transition). 16S rRNA sequence regions of ambiguous alignment, and incompletely sequenced 3' and 5' ends, were excluded from the analysis. Distances were calculated using the Kimura two-parameter model ($T = 2$) as implemented in the program DNADIST of the PHYLIP package 3.5c⁵¹, and references therein. The distance matrix was calculated from 16S rRNA positions 220–840, 846–1026, 1035–1134, 1140–1371 (*Escherichia coli* numbering), total 1,130 nucleotides. Sequence data are from refs 18–21,52.

TABLE 2 16S rRNA substitution rates of γ -subdivision sulphide-oxidizing endosymbiotic bacteria

Bivalve families with obligate endosymbionts	Symbionts of bivalve host species ¹⁸	Palaeontological age of host bivalve lineages	16S rRNA sequence distance of bivalve endosymbionts (mean $\pm$ s.d.) (%)	Time for 1% substitution of 16S rRNA (Myr)
(1) Solemyidae	<i>Solemya reidi</i> <i>Solemya velum</i>	Llanvirn, 470 Myr (refs 26,53)	7.25	64.8
(2) Lucinidae	<i>Lucina floridana</i> <i>Codakia orbicularis</i> <i>Lucinoma aequizonata</i> <i>Anodontia philippiana</i>	Middle or Lower Jurassic radiation, $\sim$ 200 Myr (refs 53,54)	4.93 ± 1.54	40.6 ± 12.7
(3) Thyasiridae	<i>Thyasira flexuosa</i>	Middle Triassic, Ladinian, 235 Myr (ref. 53)	6.24 ± 0.75	37.7 ± 4.6
(4) Vesicomyidae	<i>Calyptogena elongata</i> <i>Calyptogena magnifica</i> <i>Calyptogena</i> sp. oregon <i>Vesicomya chordata</i>	Oligocene, 35 Myr (ref. 55)	2.64 ± 0.39	13.3 ± 1.9

Rates are calibrated by comparing the palaeontological age of host bivalves with the average mutual 16S rRNA sequence distances (%) of the respective endosymbionts, as determined by Kimura two-parameter analysis ($T = 2$). 16S rRNA sequence regions of ambiguous alignment, and incompletely sequenced 3' and 5' ends were excluded from the analysis. The distance matrix was calculated from 16S rRNA positions 220–840, 846–1026, 1035–1134, 1140–1371 (*E. coli* numbering), total 1,130 nucleotides. (1) Comparison between *S. reidi* and *S. velum* symbiont; (2) comparison between endosymbionts of *Lucina floridana*, *Codakia orbicularis*, *Lucinoma aequizonata* and *Anodontia philippiana*, representing four different lucinid phylogenetic lineages, of the Jurassic radiation of the Lucinidae⁵⁴: *Lucina*, *Codakia*, *Miltha* and *Anodontia*; *Lucinoma* and *Anodontia* are Oligocene and Palaeocene genera, but derived from Jurassic ancestors, *Meso-* and *Discomiltha*, and *Jagonoma*, respectively⁵⁴; (3) comparison of the thyasirid symbiont of *Thyasira flexuosa* with the four lucinid endosymbionts of (2), time-calibrated by middle Triassic split of the Superfamily Lucinacea into the families Thyasiridae and Lucinidae⁵³. This calculation presupposes that the symbiosis was already well established within the superfamily Lucinacea at the time of the split^{18,23}; (4) comparison within the Vesicomyidae symbionts; their unusually high substitution rate could reflect either rapid genomic evolution following the recent evolutionary radiation and speciation within the group, or poor palaeontological sampling of the environmentally restricted host clams.

Mutually independent mutations of the approximately 50 different secondary-structure elements¹¹ lead to a gradual accumulation over time of substitutions in the 16S rRNA sequence, which can potentially be used as a molecular clock¹⁷. We calibrated our clock by using obligately bivalve-hosted, sulphide-oxidizing symbionts, which, like their free-living counterparts *Beggiatoa*, *Thioploca*, *Thiothrix*, *Thiomicrospira* and *Thiobacillus* (*T. caldus*, *T. ferrooxidans*, *T. thiooxidans*), represent deep-branching phylogenetic lineages emerging from near the root of the proteobacterial γ subdivision^{18–22}. Similar, but not uniform, 16S rRNA sequence distances have accumulated in these lineages, averaging $15.9 \pm 4.1\%$ substitutions per site (Table 1). Similar sequence distances demonstrate similar substitution rates between free-living and symbiotic sulphide-oxidizing bacteria, minimizing the complications imposed by large variations in rate between lineages¹¹.

To calculate substitution rates we assume congruent evolution between obligately bivalve-associated endosymbiotic bacteria and their host bivalves Solemyidae, Lucinidae, Thyasiridae and Vesicomyidae¹⁸. Supporting this association, bacterial symbiosis is argued to have provided the driving force for the evolution and radiation of the Lucinidae and Thyasiridae, as well as for their morphology and infaunal mode of life²³. A long history of symbiosis is indicated for lucinoid bivalves and *Solemya* by their association with fossil hydrothermal vents²⁴. The dependence of the Solemyidae on their endosymbionts for nutrition, the reduction or total disappearance of mouth and gut, and their conserved morphology in the fossil record, all support an early establishment of the symbiosis^{25,26}. Further, all members of the Solemyidae, Lucinidae and Vesicomyidae so far examined contain endosymbionts, arguing for symbiont-associated evolution of these families^{18,25}.

We calculated the 16S rRNA sequence distances accumulated between the different symbiont species associated with particular bivalve families, assuming that symbionts diverged in parallel with the evolution of their host bivalve family. Substitution rates are then obtained by combining these sequence distances with the palaeontological age of the respective bivalve families (Table 2). Substitution rates for endosymbionts from Solemyidae, Lucinidae

and Thyasiridae average 1% per 47.7 ± 16 Myr, or 39.1 ± 14.3 Myr including the unusually fast substitution rate of the Vesicomyidae symbionts (Table 2). These values compare to the general bacterial substitution rate of 1% per 50–60 Myr, which is based on diversified bacterial lineages and environmental or physiological calibration events¹⁷, but may not be equally applicable to all bacterial groups (such as the cyanobacteria). Our values also compare to the substitution rate of 1% per 25–50 Myr for bacterial aphid symbionts²⁷.

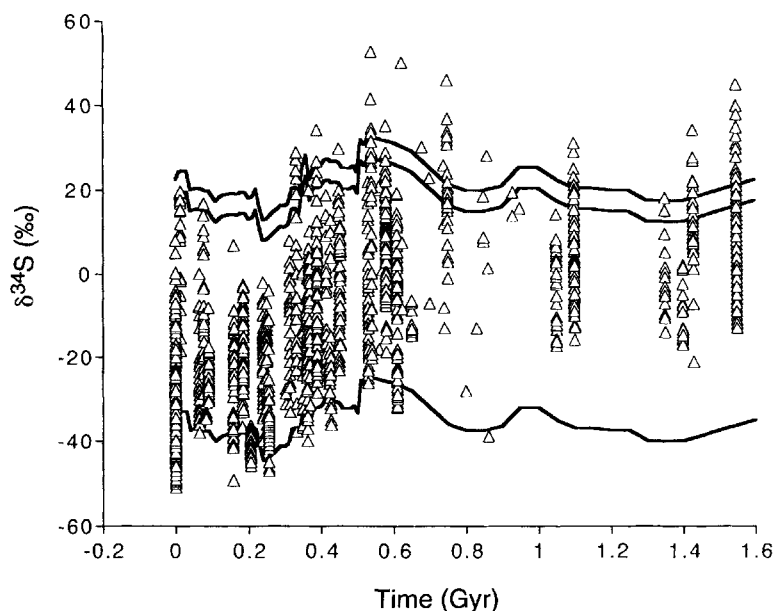
The symbiont-derived substitution rates, applied to the average interlineage distance of sulphide oxidizers ($15.9 \pm 4.1\%$), result in their evolutionary radiation at 0.76 ± 0.32 Gyr (without Vesicomyidae) or 0.62 ± 0.28 Gyr (with Vesicomyidae). Hence our data indicate a late Proterozoic origin of non-photosynthetic, sulphide-oxidizing bacteria of the γ proteobacterial subdivision. As argued previously, the requirements of these sulphide-oxidizers for oxygen reflect their evolutionary origin in a time of increased oxygenation of sulphidic bacterial habitats. Individual combinations of 16S rRNA substitutions (Table 1) and symbiont-derived substitution rates (Table 2) yield an absolute maximum age estimate of 1.44 Gyr, far younger than the 2.0 Gyr suggested for an earlier Earth-surface oxidation event^{8,28,29}.

Stable isotopes

In modern marine sediments and sulphide-rich euxinic water bodies (for example, the Black Sea), sulphide is produced by the bacterial reduction of sulphate. During this process, pure cultures of sulphate-reducing bacteria produce sulphide depleted in the isotope ^{34}S by 4–46‰ (Fig. 1), with an average of $18 \pm 10\%$ ($n = 155$). By contrast, the isotopic composition of marine sulphides exhibits much greater depletions in ^{34}S , with values in the range 24–71‰, averaging $51 \pm 10\%$ ($n = 234$) (Fig. 1). These observations reinforce previous suggestions^{30–32} that isotopic fractionation during bacterial sulphate reduction is unable to explain fully the isotopic composition of marine sulphides.

The large ^{34}S depletions of modern sedimentary sulphides are thought to result from an initial fractionation by sulphate-reducing bacteria (Fig. 1), followed by further ^{34}S -depletion during the oxidative part of the sulphur cycle^{31,32}. Results from modern, non-

FIG. 3 Isotopic composition of sedimentary sulphides from the present to the Mesoproterozoic inclusive (1.6 Gyr). Isotopic compositions from earlier times, particularly before 2 Gyr, may have been influenced by low seawater sulphate levels⁵⁷, and are not included here. The data only include sulphides with a probable association with bacterial processes. Hence sulphides from highly metamorphosed (beyond Greenschist facies) sediments, sulphides hosted in non-sedimentary rocks, and those of hydrothermal origin have not been used. Except for some copper sulphide analysis from the 1.1 Gyr Nonesuch shale (Michigan)⁵⁸, only pyrite analyses have been included. Also included is a band (with a width of 5‰) representing the isotopic composition of seawater sulphate^{59,60}, and a curve showing seawater sulphate displaced by 55‰. Seawater sulphate values are fairly certain in the Phanerozoic, becoming less certain in the Proterozoic. However, to affect the interpretations offered here, the isotopic composition of seawater sulphate would need to be consistently underestimated by 20‰ in sediments older than 1 Gyr, but there is no evidence for this. A complete list of references and data is available as Supplementary Information.



photosynthetic, siliciclastic sediments support this contention, as the depletion of sulphides in ^{34}S correlates with the extent of sulphide oxidation (Fig. 2a). Further, large ^{34}S depletions of 62‰ are observed for sulphides in the euxinic Black Sea, far exceeding the 30‰ provided by sulphate reducers from the water column of the Black Sea³³. Substantial sulphide oxidation has been demonstrated in the Black Sea³⁴. Modern photosynthetic microbial mats and marine sapropels provide an exception, with less isotopic fractionation than anticipated (Fig. 2a); these isotopic systematics will be discussed below. Finally, in contrast with previous reports^{35,36}, we found that the rate of sulphate reduction does not substantially influence the isotopic composition of sedimentary sulphides (Fig. 2b).

Elemental sulphur has been implicated as an important reaction intermediate of sulphide oxidation^{32,34,37}, with ^{34}S -depleted sulphide produced by the bacterial disproportionation of elemental sulphur to sulphate and sulphide³². Through repeated cycles of sulphide oxidation to elemental sulphur, followed by disproportionation, ^{34}S -depleted sedimentary sulphides can be generated, and the trends in Fig. 2 can be rationalized³². Other parts of the oxidative sulphur cycle, such as thiosulphate formation and disproportionation, could also lead to ^{34}S -depleted sulphides³¹, although there are as yet no reports of these isotopic systematics. In modern non-photosynthetic sediments, and euxinic waterbodies without light at the oxic/anoxic interface, sulphide is oxidized either directly or indirectly with oxygen through both biologically-mediated and inorganic pathways^{34,38} (Fig. 2). These observations form the basis for our interpretation of the isotopic record of sedimentary sulphides.

An extensive compilation of sedimentary sulphide-isotope compositions, dating from the present back to the Mesoproterozoic, is presented in Fig. 3. The data include 1,840 individual analyses, of which 1,213 are from the Phanerozoic (<0.544 Gyr) and 627 are from the Neoproterozoic and Mesoproterozoic (0.544–1.6 Gyr). Through the Phanerozoic, the most ^{34}S -depleted sulphides crudely follow the seawater sulphate curve with an approximate displacement of 55‰ (Fig. 3). Overall, 16% of the data show maximum discriminations of 55‰ or more, with 34% of all data having discriminations of equal to or greater than 45‰ compared to seawater sulphate. In only one instance has a discrimination as great as 45‰ been reported for sulphate-reducing bacteria (Fig. 1), so this cannot be taken as a typical fractionation. Hence marine sulphides are commonly formed through all of the Phanerozoic, with isotopic discriminations greater than can be generated by sulphate-reducing bacteria,

strongly indicating the operation of the oxidative sulphur cycle as described above.

A similar pattern persists into the Neoproterozoic with the deposition of the Issac Formation and the underlying Kaza Group of the Windermere Supergroup at between 0.58 Gyr and 0.64 Gyr (ref. 39). Neither the Issac Formation or the Kaza Group have any traces of burrowing organisms (G. M. Ross, personal communication). Further, although the uppermost Issac Formation contains Ediacaran fossils, this interval is well separated from the Cambrian/Precambrian boundary and the initiation of vertical burrowing and extensive sediment mixing by animals⁴⁰. Thus sulphides from the Issac Formation and Kaza Group have been formed in the absence of bioturbation and the possible influence of sediment mixing on sulphur cycling.

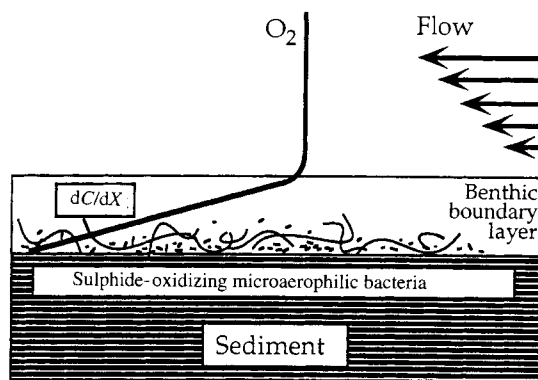
Single isolated analysis of the Little Dal Group (0.80 Gyr) and the Awatubi Member of the Chuar Group⁴¹ (0.86 Gyr) suggest that Phanerozoic patterns of fractionation may persist to 0.86 Gyr, although more detailed sampling would better indicate the persistence of these large fractionations. Before 0.86 Gyr no isotope analysis ($n = 383$) yields discriminations of $\geq 45\%$ compared to seawater sulphate, and only seven analyses yield discriminations greater than 35‰.

The isotope record thus demonstrates a fundamental change in the processes controlling the isotopic composition of sedimentary sulphides some time between 0.64 Gyr and 1.05 Gyr. This is consistent with the wide-scale initiation of the oxidative sulphur cycle, as driven by the accumulation of oxygen in the atmosphere and marine environments. This time frame is contemporaneous with the evolution of sulphide-oxidizing bacteria, and the liberation of oxidants during the enhanced burial of organic carbon during the late Proterozoic¹⁰. Although this pre-dates the appearance of the Ediacaran fauna, it is consistent with the evolution of the first metazoans, as indicated by amino-acid comparisons⁵.

Transition from low to higher oxygen levels

Before the proposed oxygenation of marine environments and the evolution of oxygen- and nitrate-respiring non-photosynthetic sulphur bacteria, sulphide oxidation would probably have occurred by anoxygenic phototrophic bacteria (representing, in part, very ancient phylogenetic lineages¹¹) in the water column, and, on the sea floor, in photosynthetic oxygen-producing, microbial mats and stromatolites. Benthic photosynthetic communities are well known in the middle and late Proterozoic⁴², and low oxygen concentrations would favour the widespread occurrence of ocean anoxia and associated sulphidic conditions⁴³. The isotope

FIG. 4 Schematic representation of the model used to constrain the water-column oxygen levels needed to oxygenate surface sediments. The supply of oxygen to the sediment surface is limited by the concentration of oxygen in the overlying water and the thickness of the benthic boundary layer by Fick's first law: $O_2 \text{ flux} = D_{O_2} (dC/dx)$ where $O_2 \text{ flux}$ is the flux of oxygen to the sediment surface, D_{O_2} is the diffusion coefficient of oxygen and dC/dx is the vertical concentration gradient in the benthic boundary layer (BBL) with C concentration and x distance. With C_w the concentration of oxygen in the overlying water, L the thickness of the BBL and the assumption that O_2 goes to zero at the sediment surface, $dC/dx = C_w/L$. The parameters D_{O_2} and L are fixed for a given temperature and benthic boundary layer thickness. The $O_2 \text{ flux}$ is fixed by the demand for oxygen as determined by the sediment metabolic rate (see text). The value of C is thus uniquely provided for given values of D_{O_2} , L and $O_2 \text{ flux}$. This calculation assumes diffusion through a benthic boundary layer 500 μm thick⁶¹, and a diffusion coefficient for oxygen at 15 °C of $1.72 \times 10^{-5} \text{ cm}^2 \text{ s}^{-1}$. A complete list of references and data is available as Supplementary Information.



record suggests that sulphur processing under such circumstances should not have led to large depletions of sulphide in ^{34}S (Fig. 3). We review the circumstances that may have led to reduced fractionations.

In photosynthetic systems with low sulphide availability, anoxygenic photosynthetic bacteria generally oxidize sulphide directly to sulphate⁴⁴, with only small associated fractionations⁴⁵. Thus the formation of elemental sulphur or other sulphur intermediates, and the associated fractionations during their disproportionation, are not favoured. When sulphide availability is high, elemental sulphur is a preferred oxidation intermediate, which, with all *Chlorobium* and *Ectothiorhodospira* spp., some Rhodospirillaceae, and some *Chloroflexus* spp., is excreted from the cell to the environment^{44,46}. This elemental sulphur may be oxidized further by anoxygenic phototrophs to sulphate. The utilization of elemental sulphur by disproportionating bacteria is hindered by the lack of growth of these organisms at even relatively moderate sulphide concentrations of $>1 \text{ mM}$ (ref. 47). Thus smaller ^{34}S sulphide depletions in sedimentary sulphides before 0.86 Gyr could be explained by more direct oxidation of sulphide to sulphate by anoxygenic phototrophs, or an inhibition of the growth of elemental sulphur-disproportionating bacteria as a result of high sulphide levels.

The reduced ^{34}S -sulphide depletions of modern microbial mats and photosynthetic sapropels (Fig. 2a) can be explained by a similar combination of factors. Mats support rapid rates of sulphate reduction, which frequently lead to high ($>1 \text{ mM}$) concentrations of dissolved sulphide⁴⁸. Further, anoxygenic photosynthetic bacteria are common members of modern mat communities⁴⁸. Finally, the daytime accumulation of oxygen from oxygenic photosynthesis may limit the growth of strictly anaerobic elemental sulphur-disproportionating bacteria. Except for the late Proterozoic introduction of non-photosynthetic sulphide oxidizers, these factors have probably changed little over the course of time considered here, explaining why the sulphur systematics in mats have been little affected by the late Proterozoic rise in oxygen.

By contrast, the invasion of oxygen would have profoundly affected euxinic water bodies. Possible changes include a deepening of the oxic/anoxic interface, decreased sulphide levels owing to increased aerobic carbon oxidation, the direct reaction of sulphide with oxygen, producing more complex sulphide intermediates, and, after their evolution, the colonization of, and increasing dominance of, non-photosynthetic sulphide oxidizers. Taken together, these factors probably contribute to the production and disproportionation of sulphur intermediates, producing increased ^{34}S sulphide depletions, such as those seen in the Black Sea³³ and sulphate-rich meromictic lakes like Mahoney Lake, British Columbia⁴⁹.

Even though some euxinic water bodies, such as Mahoney Lake, support active anoxygenic photosynthesis, fully aerobic upper layers produce interfaces supporting steep opposing gradients of

oxygen and sulphide and provide the possibility of both direct chemical and non-photosynthetic bacterial sulphide oxidation, particularly during low-light winter months. Hence high oxygen levels create chemical circumstances quite different than envisioned for middle Proterozoic oceans, limiting the usefulness of modern euxinic water bodies as direct analogues for ancient oceans.

A further consequence of increased oxygen is its diffusion into non-photosynthetic marine sediments. With insufficient oxygen, the sulphide produced by sulphate reduction would diffuse from the sediment, for probable oxidation by phototrophs in the water column, as described above. The penetration of oxygen into the sediment allows multiple respiratory pathways of carbon oxidation (other than just sulphate reduction), and the complex redox cycling of modern sediments³⁸. These same circumstances are thought to initiate the oxidative sulphur cycle, producing highly ^{34}S -depleted sulphides.

Constraint on oxygen levels

There is a general requirement of *Beggiatoa* to live bound to sediment in steep opposing gradients of sulphide and oxygen¹⁴. Some representatives of other non-photosynthetic sulphide-oxidizing genera are also bound to sediment and require oxygen (see above). Hence the evolution of these sulphide oxidizers, as with large ^{34}S sulphide depletions in non-photosynthetic marine sediments, should be contemporaneous with oxygenation of the sediment surface. We offer an estimate of the oxygen levels required to oxygenate non-photosynthetic marine sediments.

Oxygen diffuses to the sediment surface along a concentration gradient through a stagnant benthic boundary layer (Fig. 4). To penetrate to and through the sediment surface, the flux of oxygen must meet or exceed the demands for oxygen in the sediment. Sediment oxygen demand is determined by the carbon oxidation rate of the sediment, regardless of whether oxygen is used to oxidize carbon directly, or the reduced products of anaerobic carbon oxidation including sulphide, Fe^{2+} , Mn^{2+} and ammonia³⁸.

A likely range of oxygen demand is derived from an extensive compilation of carbon oxidation rates for modern coastal sediments in water $<200 \text{ m}$ deep, a depth likely to maintain approximate exchange equilibrium with atmospheric oxygen. In a relatively low-oxygen atmosphere, which will satisfy the minimum oxygen requirements of this calculation, greater water depths may have been completely anaerobic⁴³ or have experienced significant oxygen depletions, and are thus not of interest here. Carbon oxidation rates range between 3.9 and 82 $\text{mmol m}^{-2} \text{ d}^{-1}$, with a median of 13.7 $\text{mmol m}^{-2} \text{ d}^{-1}$ ($n = 60$). Our assumption that late Proterozoic coastal sediments experienced a similar range of rates is supported by a broad similarity between the average organic carbon contents of modern (0.65 wt%) and late Proterozoic sediments (0.4 wt%)⁵⁰, accounting for some carbon loss through deep burial and thermal maturation. The carbon oxidation rates given above correspond to a similar range of

oxygen demand, assuming that one mole of oxygen oxidizes one mole of organic carbon. Using Fick's first law, this range of oxygen demand is balanced by diffusion from overlying water containing oxygen at between 13 and 552 μM , with a median value of 46 μM (Fig. 4).

Only a very small proportion of sediments would have been exposed to oxygen with a water-column concentration of 13 μM , which translates into 5% PAL with an air saturated PAL of 254 μM at 15 °C. A higher proportion of oxic sediments than accounted for by this minimum oxygen level might be required to induce an evolutionary radiation of sulphide-oxidizing bacteria, and to uncover highly ^{34}S -depleted sulphides through random sampling. Perhaps somewhat arbitrarily, although not unreasonably, we suggest that bacterial evolution would have been induced, and sufficient aerial coverage for highly ^{34}S -depleted sulphides provided, by the time that half of the coastal sea floor was aerobic. This would occur at a bottom water oxygen level of 46 μM ,

requiring 18% PAL. Thus, to induce an evolutionary radiation of sulphide-oxidizing bacteria, and to explain the trend in sedimentary sulphur isotopes, our best estimate is that, between 0.64 Gyr and 1.05 Gyr, atmospheric oxygen passed from lower than to higher than 5–18% PAL.

This rise in oxygen levels considerably post-dates an earlier Earth surface oxidation event at around 2.0 Gyr (refs 8, 28, 29), for which the modelling of soils developed on basalts²⁸ yields estimates of atmospheric oxygen of greater than or equal to 1% PAL. The retention of iron during the weathering of the Kuruman Iron Formation from South Africa, at 2.2–1.9 Gyr, suggests higher levels of up to 15% or greater PAL²⁹. Our results suggest that a maximum oxygen concentration of 5–18% PAL was attained during this oxidation event, and that higher levels were not reached until the late Proterozoic, allowing for a fully operative oxidative sulphur cycle, the evolution of non-photosynthetic sulphide-oxidizing bacteria, and, ultimately, metazoan life. $\square$

Received 22 January; accepted 5 June 1996.

- Runnegar, B. *Palaeontology* **29**, 1–24 (1986).
- Glaessner, M. F. *Geol. Soc. Am. Bull.* **82**, 509–514 (1971).
- Bowring, S. A. et al. *Science* **261**, 1293–1298 (1993).
- Berkner, L. V. & Marshall, L. C. *J. Atmos. Res.* **22**, 225–261 (1965).
- Runnegar, B. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **97**, 97–111 (1991).
- Knoll, A. H. in *Origin and Early Evolution of the Metazoa* (eds Lipps, J. H. & Signor, D. W.) 53–84 (Plenum, New York, 1992).
- Towe, K. M. *Proc. natn. Acad. Sci. U.S.A.* **65**, 781–788 (1970).
- Des Marais, D. J., Strauss, H., Summons, R. E. & Hayes, J. M. *Nature* **359**, 605–609 (1992).
- Holland, H. D. *The Chemistry of the Atmosphere and Oceans* (Wiley, New York, 1978).
- Knoll, A. H., Hayes, J. M., Kaufman, A. J., Swett, K. & Lambert, I. B. *Nature* **321**, 832–838 (1986).
- Woese, C. R. *Microbiol. Rev.* **51**, 221–271 (1987).
- Fox, G. E. et al. *Science* **209**, 457–463 (1980).
- Bland, J. A. & Staley, J. T. *Arch. Microbiol.* **117**, 79–87 (1978).
- Jørgensen, B. B. *Phil. Trans. R. Soc. Lond. B* **298**, 543–561 (1982).
- Fossing, H. et al. *Nature* **374**, 713–715 (1995).
- Kuenen, J. G., Robertson, L. A. & Tuovinen, O. H. in *The Prokaryotes* (eds Balows, A., Trüper, H. G., Dworkin, M., Harder, W. & Schleifer, K.-H.) 2638–2657 (Springer, New York, 1991).
- Ochman, H. & Wilson, A. C. *J. molec. Evol.* **26**, 74–86 (1987).
- Distel, D. L., Felbeck, H. & Cavanaugh, C. M. *J. molec. Evol.* **38**, 533–542 (1994).
- Muyzer, G., Teske, A., Wirsén, C. O. & Jannasch, H. W. *Arch. Microbiol.* **164**, 165–172 (1995).
- Polz, M. F., Odintsova, E. & Cavanaugh, C. M. *Int. J. Syst. Bact.* **46**, 94–97 (1996).
- Teske, A., Ramsing, N. B., Küver, J. & Fossing, H. *Syst. appl. Microbiol.* **18**, 517–526 (1995).
- Lane, D. J. et al. *J. Bact.* **174**, 269–278 (1992).
- Reid, R. G. B. & Brand, D. G. *Veliger* **29**, 3–24 (1986).
- Campbell, K. A. & Bottjer, D. J. *Natn. geogr. Res. Explor.* **9**, 326–343 (1993).
- Cavanaugh, C. M. *Am. Zool.* **34**, 79–89 (1994).
- Pojan, J. J. *New Mex. Bur. Mines miner. Resour. Mem.* **44**, 201–271 (1988).
- Moran, N. A., Munson, M. A., Baumann, P. & Ishikawa, H. *Proc. R. Soc. Lond. B* **253**, 167–171 (1993).
- Holland, H. D., Feakes, C. R. & Zbinden, E. A. *Am. J. Sci.* **289**, 362–389 (1989).
- Holland, H. D. & Beukes, N. J. *Am. J. Sci.* **290**, 1–34 (1990).
- Kemp, A. L. W. & Thode, H. G. *Geochim. cosmochim. Acta* **32**, 71–91 (1968).
- Jørgensen, B. B. *Science* **249**, 152–154 (1990).
- Canfield, D. E. & Thamdrup, B. *Science* **266**, 1973–1975 (1994).
- Fry, B. et al. *Deep-Sea Res.* **38**, S655–S661 (1991).
- Jørgensen, B. B., Fossing, H., Wirsén, C. O. & Jannasch, H. W. *Deep-Sea Res.* **38**, S1083–S1103 (1991).
- Goldhaber, M. B. & Kaplan, I. R. *Soil Sci.* **119**, 42–55 (1975).
- Ohmoto, H., Kakegawa, T. & Lowe, D. R. *Science* **262**, 555–557 (1993).
- Pyzik, A. J. & Sommer, S. S. *Geochim. cosmochim. Acta* **45**, 687–698 (1981).
- Canfield, D. E. et al. *Mar. Geol.* **113**, 27–40 (1993).
- Ross, G. M., Bloch, J. D. & Krouse, H. R. *Precamb. Res.* **73**, 71–99 (1995).
- Fedonkin, M. A. & Runnegar, B. N. in *The Proterozoic Biosphere* (eds Schopf, J. W. & Klein, C.) 389–395 (Cambridge Univ. Press, 1992).
- Hayes, J. M., Lambert, I. B. & Strauss, H. in *The Proterozoic Biosphere* (eds Schopf, J. W. & Klein, C.) 129–132 (Cambridge Univ. Press, 1992).
- Sergeev, V. N., Knoll, A. H. & Grotzinger, J. P. *J. Paleont.* **69**, 1–37 (1995).
- Logan, G. A., Hayes, J. M., Hieshima, G. B. & Summons, R. E. *Nature* **376**, 53–56 (1995).
- Pfennig, N. *Pl. Soil* **43**, 1–16 (1975).
- Chambers, L. A. & Trudinger, P. A. *Geomicrobiol. J.* **1**, 249–293 (1979).
- Trüper, H. G. & Fischer, U. *Phil. Trans. R. Soc. Lond. B* **298**, 529–542 (1982).
- Thamdrup, B., Finster, K., Hansen, J. W. & Bak, F. *Appl. env. Microbiol.* **59**, 101–108 (1993).
- Pierson, B. K. et al. in *The Proterozoic Biosphere* (eds Schopf, J. W. & Klein, C.) 247–340 (Cambridge Univ. Press, 1992).
- Overmann, J., Beatty, J. T., Krouse, H. R. & Hall, K. J. *Limnol. Oceanogr.* **41**, 147–156 (1996).
- Schidlowski, M. in *Organic Geochemistry* (eds Engel, M. H. & Macko, S. A.) 639–655 (Plenum, New York, 1993).
- Felsenstein, J. *PHYLP (Phylogeny Inference Package)* (Seattle, 1993).
- Maidak, B. L. et al. *Nucleic Acids Res.* **22** (suppl.), 3485–3487 (1994).
- Cox, D. L. et al. in *Treatise on Invertebrate Paleontology* (eds Moore, R. C.) N225–N868 (Geological Society of America, Lawrence, KS, 1969).
- Bretsky, S. S. in *Palaeontogr. am.* **8**, 217–338 (1976).
- Goedert, J. L. & Squires, R. L. *Veliger* **36**, 72–77 (1993).
- Nakai, N. & Jensen, M. L. *Geochim. cosmochim. Acta* **28**, 1893–1912 (1964).
- Cameron, E. M. *Nature* **296**, 145–148 (1982).
- Burnie, S. W., Schwarcz, H. P. & Crockett, J. H. *Econ. Geol.* **67**, 895–914 (1972).
- Lind, T. B. thesis, Univ. Miami (1983).
- Strauss, H. *Precamb. Res.* **63**, 225–246 (1993).
- Gundersen, J. K. & Jørgensen, B. B. *Nature* **345**, 604–607 (1990).

ACKNOWLEDGEMENTS. We thank D. Des Marais, A. Knoll and B. B. Jørgensen for comments, and B. Thamdrup and G. Ross for helpful discussions. This work was supported by NASA and the Max Planck Society.

CORRESPONDENCE should be addressed to D.E.C.

KNOW YOUR COPY RIGHTS RESPECT OURS

The publication you are reading is protected by copyright law.
Photocopying copyright material without permission is no different from stealing a magazine from a newsagent, only it doesn't seem like theft.

If you take photocopies from books, magazines and periodicals at work your employer should be licensed with CLA.

Make sure you are protected by a photocopying licence.



The Copyright Licensing Agency Limited
90 Tottenham Court Road, London W1P 0LP
Telephone: 0171 436 5931 Fax: 0171 436 3986

Integration of positional signals and regulation of wing formation and identity by *Drosophila vestigial* gene

Jaeseob Kim, Angela Sebring, Jeffrey J. Esch, Mary Ellen Kraus, Kathy Vorwerk, Jeffrey Magee & Sean B. Carroll

Howard Hughes Medical Institute and Laboratory of Molecular Biology, University of Wisconsin, 1525 Linden Drive, Madison, Wisconsin 53706, USA

Appendage formation is organized by signals from discrete sources that presumably act upon downstream genes to control growth and patterning. The *Drosophila vestigial* gene is selectively required for wing-cell proliferation, and is sufficient to induce outgrowths of wing tissue from eyes, legs and antennae. Different signals activate separate enhancers to control *vestigial* expression: first, in the dorsal/ventral organizer through the *Notch* pathway, and subsequently, in the developing wing blade by *decapentaplegic* and a signal from the dorsal/ventral organizer. Signal integration must be a general feature of genes like *vestigial*, that regulate growth or patterning along more than one axis.

ARTHROPOD and vertebrate appendages develop as outgrowths of the embryonic body wall after the primary anteroposterior (A/P) and dorsoventral (D/V) axes have been established. The growth and patterning of these structures requires the integration of multiple signals emanating from discrete sources within the imaginal disc or limb bud. In developing vertebrate limbs, the zone of polarizing activity (ZPA) and the apical ectodermal ridge (AER) serve as signalling sources by producing the *Sonic hedgehog* (*Shh*) and *fibroblast growth factor* (*FGF8* and *FGF4*) proteins, respectively (see ref. 1 for review). *Shh* regulates A/P polarity² within the mesenchymal layer, whereas the *FGF* family members³ are necessary for proximodistal (P/D) axis formation. These two signals and *Wnt-7a*, which organizes D/V polarity⁴⁻⁶, act in concert to control limb growth and patterning⁷⁻⁹.

In the *Drosophila* wing imaginal disc, the A/P and D/V compartment boundaries are discrete signalling sources (see ref. 10 for review). Cells in the posterior compartment express the *engrailed* (*en*) selector gene which activates expression of the signalling molecule gene, *hedgehog* (*hh*)^{11,12}. The *hh* protein then diffuses into the anterior compartment, where cells along the A/P boundary respond to the *hh* signal¹³⁻¹⁵ and produce a second signalling molecule, the *decapentaplegic* protein Dpp. Dpp then diffuses bidirectionally to organize growth and patterning of the wing disc along the A/P axis¹⁶⁻¹⁹. The D/V compartmentalization of *Drosophila* wing discs is regulated by *apterous* (*ap*), a LIM (for *lin-11 isl-1 mec-3*) homeodomain selector gene²⁰⁻²², which is expressed in all dorsal cells. The *ap* protein activates the expression of *fringe* (*fng*)²³, a putative extracellular protein that functions as a boundary recognition molecule. When *fng*-expressing dorsal cells are juxtaposed with ventral cells that do not express *fng*, expression of *Serrate* (*Ser*)²⁴, which produces a ligand for the *Notch* (*N*) receptor²⁵, is upregulated²⁶. *Ser* is necessary for both wing formation^{24,27} and the patterning of the wing edge, or margin²⁶⁻²⁹.

All of these vertebrate and *Drosophila* signals are presumed to act by regulating the expression of other signals. In *Drosophila*, the loss of either A/P or D/V selector-gene or signal functions abolishes wing development, suggesting that regulatory inputs from both patterning axes are required for the formation of appendage structures. As the *vestigial* protein is required for wing formation, expressed in the region of the wing disc that will form the wing proper, and activated along the dorsoventral boundary by a *Ser*-dependent mechanism^{26,27}, the *vestigial* gene

may be a crucial target for the organizing signals from the compartment boundaries. Here, we show that *vestigial* is selectively required for the proliferation of wing cells, and that targeted expression of *vestigial* is sufficient to induce wing-like outgrowths from many other imaginal tissues. The regulation of *vestigial* expression in discrete sectors of the wing is controlled by separate *cis*-acting regulatory elements that respond to different signalling proteins. Expression of *vg* at the D/V boundary is regulated directly by the *ap-fng-Ser-N* pathway by the *Suppressor of Hairless* (*Su(H)*) protein through one enhancer, whereas *vg* expression in developing wing-blade cells requires the *dpp* signal from the A/P boundary and a signal from the D/V boundary which act through a second enhancer. The activation of the *vestigial* gene is thus a 'nodal point' that connects compartmentalization to the control of wing formation and identity.

***vg* is selectively required in wing cells**

To examine further the requirement for *vestigial* function during wing formation, we have analysed the behaviour of clones of cells that lack *vestigial* activity. Because even partial loss-of-function *vg* mutant clones are difficult to recover in the adult wing³⁰, we concentrated on the behaviour of *vg* null mutant clones in the developing wing disc. There is a striking difference in *vg*⁻ clone behaviour between the developing wing primordia (where *vestigial* is normally expressed; Fig. 1a) and the presumptive notum (where it is not). Large *vg*⁻/*vg*⁻ clones, comparable in size to the *vg*⁺/*vg*⁺ twin spots, were readily recovered in the notum (Fig. 1a, b), whereas only tiny *vg*⁻/*vg*⁻ clones of two, or at most three, cells could be found in the developing wing pouch (Fig. 1c-e). As development proceeds, *vg*⁻ clones in the wing pouch do not proliferate; older large twin spots are accompanied by no or few *vg*⁻ cells (Fig. 1c-e). This result suggests that the *vestigial* protein is selectively required in the presumptive wing for cell proliferation.

***vestigial* regulates wing identity**

The absolute requirement for *vestigial* in wing cells suggested that this protein could be involved in wing-cell type-specific differentiation. We explored this possibility by ectopically expressing the *vestigial* protein in other imaginal discs using the GAL4/UAS targeted gene expression system³¹. We placed a *vg* complementary DNA under the control of a yeast UAS regulatory element (*UAS-vg*) and crossed *UAS-vg* flies to many different *GAL4*

insertion lines. Four crosses produced offspring with pharate adult phenotypes (many others were lethal at the early pupal stage). Ectopic wing-like outgrowths were formed from the eyes, head capsule, forelegs and antennae (Fig. 2a–c), in positions consistent with the expression patterns of the *GAL4* lines as determined with

a *UAS-lacZ* detector construct. Outgrowths replace normal tissue with an excess of wing-like tissue, indicating excess proliferation as well as transdetermination. Imaginal disc overgrowth was apparent in the third instar and restricted to *vg*-expressing cells as revealed by *UAS-lacZ* expression (Fig. 2h, i and data not shown).

FIG. 1 *vg* is selectively required in the wing primordia for cell proliferation. a, In the wing primordia (w) a double intensity Myc-labelled mitotic twinspace (Myc shown in green) is visible but an accompanying *vg*[−] clone (*vg* protein in red) is not. A large *vg*[−] clone (detected by loss of Myc staining), and accompanying mitotic twinspace are evident in the notal region of the wing disc (n). As the *vg*[−] clone in the notum and the clone absent from the wing would be of the same genotype, the difference in survival depends upon their position in the disc. Clones (arrows) and accompanying mitotic twinspace (arrowheads) are indicated. Magnification 15×. b–e, High magnification views (60×) of b, the notal clone in a; c, the wing mitotic twinspace in a; d, e, a three-cell *vg*[−] clone and twinspace in the wing pouch. The *vg*[−] clone is outlined. These *vg*[−] clones do not proliferate further.

METHODS. A *vg*^{83b27R/Gla}, *Bc* (ref. 24) stock was mated with a heat-shock inducible 42 μ M *myc* stock²⁶. 72–84-h staged larvae were collected and mitotic recombination was induced using γ -rays²⁶. After 60 h (a–c) or 36 h (d, e) of growth at 25 °C, larvae were heat shocked for 2 h at 37 °C to induce *myc* expression. After a 90-min recovery at 25 °C, wing imaginal discs from non-*Bc* larvae were fixed and stained²⁰ with rabbit anti-vestigial²⁴ and mouse anti-Myc antibodies²⁶.

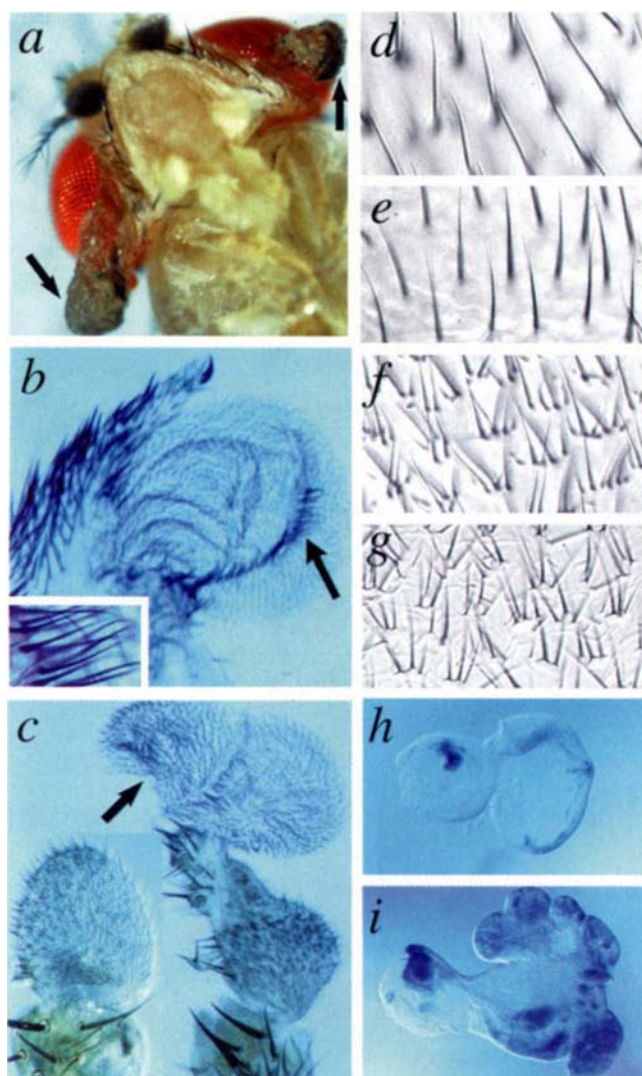
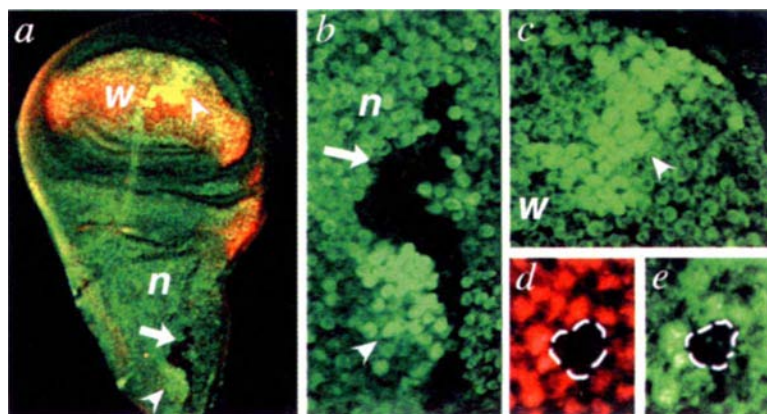
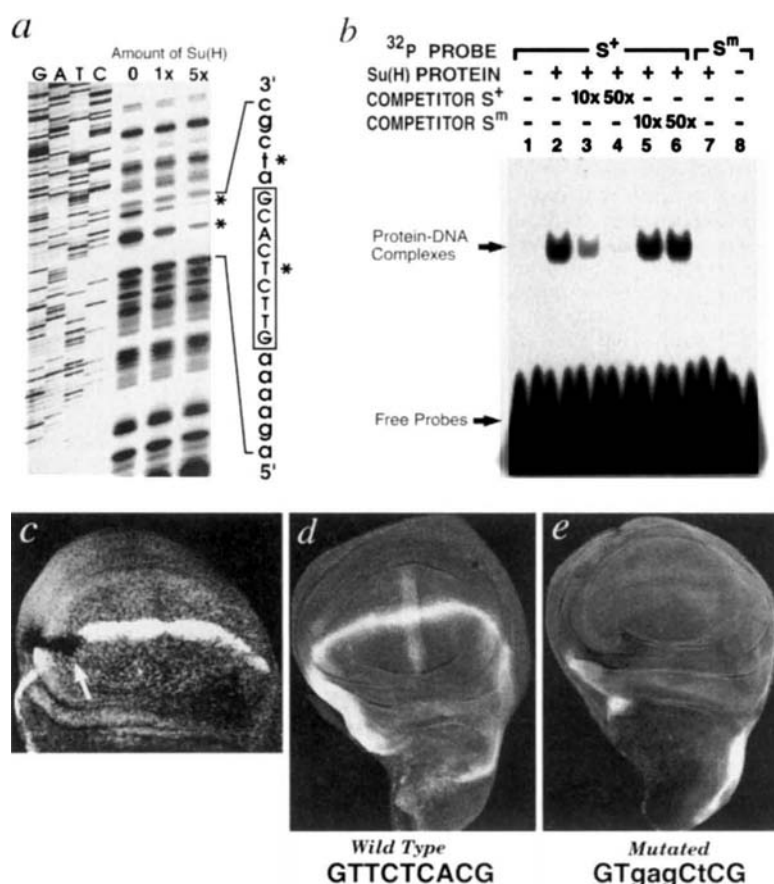


FIG. 2 Targeted *vestigial* expression induces wing-like outgrowths. a, adult *UAS-vg/+*; *GAL4-dpp4A.3/+* fly, ventral view, with wing like outgrowths from the ventral portion of both eyes (arrows). b, *GAL4-E132*; *UAS-vg/+* second thoracic leg, with a wing-like outgrowth from the base of the tibia (16×). Inset: high-magnification (100×), showing wing-like hairs and margin bristles. c, Segments 2 and 3 of a wild-type antenna (left) and an antenna with a *vestigial*-induced outgrowth (right) (*UAS-vg*; *GAL4-E132*; ref. 50). Magnification is the same for both antennae (16×). d, Wild-type wing-blade hairs. e, Hairs from a head-capsule outgrowth (*UAS-vg*; *GAL4#409*). Note the similar morphology and spacing to wild-type wing-blade hairs. f, Wing-blade hairs, *multiple wing hairs* phenotype. g, Head capsule outgrowth from a *UAS-vg/+*; *mwh/mwh* *GAL4-dpp4A.3* adult fly, showing the wing-specific *mwh* phenotype. h, Eye antennal disc, *UAS-lacZ*; *GAL4-dpp4A.3* genotype, stained with X-gal. i, Antennal disc of the genotype *UAS-lacZ/UAS-vg*; *GAL4-dpp4A.3/+*. The *lacZ*-expressing region (blue) has expanded, and the disc is overgrown. Magnification is the same as in h (16×).

METHODS. The *vestigial* open reading frame was amplified by polymerase chain reaction (PCR) from *Drosophila* cDNA-3 (ref. 24), inserted into pUAST (ref. 26), and injected into *Drosophila*. *UAS-vg* females were crossed with males (*GAL4-dpp4A.3* (ref. 37), *GAL4-E132* (ref. 1), or *GAL4#409*; D. Brower, unpublished results) and the progeny grown at 25 °C until the pharate stage; tissues in b–g were mounted in Faure's mount. For g, *mwh* was recombined with *GAL4-dpp4A.3* and *UAS-vg*; these two resultant lines were crossed, giving *UAS-vg/+*; *mwh/mwh* *GAL4-dpp4A.3* genotype flies. h, i, Female *UAS-lacZ* or *UAS-lacZ*; *UAS-vg* flies were crossed with male *GAL4-dpp4A.3/TM6* flies and the progeny grown at 25 °C until the late third instar. Non-*TM6* larvae were dissected in phosphate-buffered saline and stained with X-gal (ref. 38).

FIG. 3 Su(H) protein binds to the vg D/V boundary enhancer *in vitro* and the Su(H)-binding site is essential for vg enhancer activation *in vivo*. **a**, DNase I footprint of Su(H) protein on vg D/V-boundary enhancer (last three lanes) with the DNA sequencing products of the vg D/V-boundary enhancer (first four lanes, marked with their corresponding ddNTP reaction). The relative amounts of Su(H) protein added to the binding reaction are indicated: 0, for no Su(H) protein; $\times 1$ for an equimolar amount of Su(H) protein and probe DNA; $\times 5$, for a fivefold molar excess of Su(H) protein compared with the probe DNA. The two DNase I-sensitive bases protected by Su(H) proteins are marked with asterisks. The core Su(H) recognition site is boxed. **b**, Competition gel-mobility shift assay for the binding specificity of Su(H) protein on the vg D/V-enhancer DNA. The addition of each component is indicated by + or – or 10 $\times$ and 50 $\times$ for the relative molar excess. The wild-type vg D/V enhancer DNA fragment probe (labelled with 32 P) forms shifted complexes well (lane 2), whereas the vg enhancer with an altered Su(H) binding site does not (lane 7). Double-stranded oligonucleotides containing the wild-type Su(H) binding site compete efficiently with the vg D/V-enhancer DNA for the Su(H) binding (lanes 3 and 4) whereas the oligonucleotides with the mutated Su(H) site do not (lanes 5 and 6). **c**, Su(H) null clone (blank area, arrow) crossing the D/V boundary, lacks vg D/V boundary enhancer expression. **d**, The expression pattern of the wild-type vg D/V-boundary enhancer driving a *lacZ* reporter gene. **e**, The expression pattern of the vg D/V enhancer with the mutated Su(H) site. Note the D/V stripe is abolished but the notal staining remains. The mutated bases are indicated in lower case beneath. The stripe on the A/P boundary is also eliminated, suggesting that this stripe may be Su(H) regulated.

METHODS. Probes for the DNase I footprinting and the gel-mobility shift assay were synthesized by PCR. The competitor S⁺ sequence is ACAGAAAGTCTCAGATCGCTGG. The competitor S^m sequence is ACAGAAAGTgagCtCGATCGCTGG (mutated bases are shown in lower case).



Adult wings of various *GAL4/UAS-vg* genotypes were somewhat disorganized; they appeared reduced in size, and blistered or ballooned, although at the cellular level, wing identity was maintained. We attribute the abnormal wing morphology to the abnormal timing of *vestigial* expression in selected *GAL4* lines, which may disrupt the relative timing of morphogenetic processes.

Most ectopic outgrowths show no evidence of the wing patterning elements that depend upon anterior-posterior and dorsal-ventral polarity, such as wing veins and wing margin, respectively (except for those on selected legs, Fig. 2b). Thus, the outgrowths are not complete wings. However, they do resemble wing tissue in that the cells are flat and bear hairs which resemble morphologically those found on the wing blade (Fig. 2d, e). Because of the resemblance of these hairs to those of the wing blade, we tested whether these cells were of wing identity using the *multiple wing hairs* (*mwh*) marker³². *mwh* flies display a characteristic tufting of the hairs on the wing blade, with between two and seven smaller hairs in place of each normal hair³³ (Fig. 2f). The hairs on the other parts of the body are disorganized and increased in number, but they are not tufted. In a *mwh* background, the *vestigial*-induced outgrowths display the wing-specific tufting phenotype, indicating that they are indeed wing tissue (Fig. 2g).

If *vestigial* is necessary and sufficient for wing-cell growth and differentiation, then we would expect that wing formation could be induced in mutants lacking wings. Specifically, if *vestigial* is regulated by compartmental signalling sources, it should be possible to rescue wing formation in mutants lacking these sources. The dorsally restricted *apterous* selector gene plays a central role in organizing wing formation by establishing the D/V boundary^{20,22} and regulating the expression of the *fringe*²³ and *Serrate*²⁷ signals. In *ap*⁻ flies, which normally lack wings, targeted expression of *UAS-vg* along the A/P boundary using a *dpp-GAL4* driver restores wing outgrowth (data not shown). This suggests that in normal wing development, compartmentally regulated

signalling proteins operate through downstream effector genes to promote wing growth and the differentiation of pattern elements. To examine how *vestigial* is regulated by upstream signals, we have undertaken a detailed analysis of the *cis*-regulatory elements and *trans*-acting factors that control *vg* expression.

Su(H) activates the vg D/V boundary enhancer

We have previously shown that the symmetrical expression of *vg* on both sides of the D/V boundary occurs through the *N* pathway²⁶. However, the molecular mechanism of *vg* induction is not known. As *Ser* produces a ligand for the *N* receptor²⁵ and the *Su(H)* protein has been shown to activate target gene transcription in response to N-signal reception³⁴⁻³⁶, we examined the requirement for *Su(H)* function at the D/V compartment boundary. We generated patches of *Su(H)*⁻ cells in wing discs using a mitotic clonal analysis technique³⁷ and examined gene expression within and around these clones. The *Su(H)*⁻ patches that span the D/V boundary fail to activate the *vg* D/V boundary enhancer³⁸ in cells on both sides of the D/V boundary (Fig. 3c).

We tested whether *Su(H)* regulation of the *vg* boundary enhancer was direct or mediated indirectly through a different target gene. DNase I foot-printing with *Su(H)* protein on the 750-base-pair (bp) *vg* D/V enhancer revealed the presence of one strong *Su(H)* binding site (Fig. 3a). The sequence of this site is identical to one of the *Su(H)* binding sites in the *E(spl)m8* promoter^{35,36}. The binding is very specific; only wild-type oligonucleotide competitors compete efficiently for *Su(H)*-binding whereas mutated competitors do not (Fig. 3b).

To determine whether the *Su(H)* binding site is essential for activation of the *vg* D/V boundary enhancer *in vivo*, we generated transgenic fly lines in which four of nine bases of the *Su(H)* site in the *vg* D/V boundary enhancer were altered by substituting the core *Su(H)*-recognition sequence with a *SacI* restriction site. This mutant transgene is not activated along the D/V boundary in the

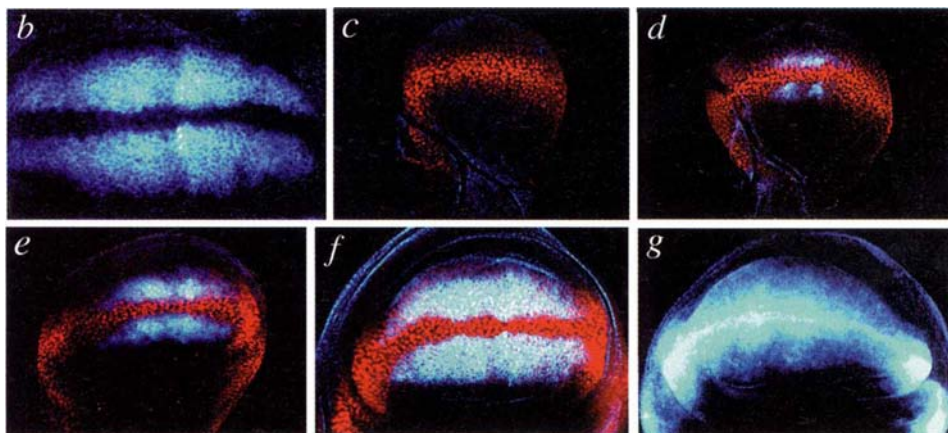
FIG. 4 The *vg*-quadrant enhancer regulates *vg* expression in the developing wing blade after the *vg* D/V-boundary enhancer is activated. In all panels, native *vg* protein is shown in red and β -galactosidase expression driven by the *vg* quadrant enhancer is shown in blue. The ventral region is at the top. **a**, DNA sequence of the 806-bp *vg*-quadrant enhancer. A potential E(spl) binding site is boxed. **b**, β -galactosidase expression driven by the *vg*-quadrant enhancer in mid-late third instar wing discs is in all cells but those at the D/V boundary. **c–f**, Temporal expression pattern of quadrant enhancer from early third to mid-late third instar stage. By the early third instar (**c**), the quadrant enhancer is not yet activated and only the *vg* D/V-boundary enhancer drives *vg* expression in cells straddling the D/V boundary. Later in the early third instar stage, the *vg*-quadrant enhancer is activated around the intersection of the A/P and D/V boundaries (**d**) and gradually fills in the entire wing pouch (**e**, **f**), whereas the *vg* D/V-boundary enhancer is being refined to a four-cell-wide sharp stripe²⁴. **g**, Placing the *vg*-quadrant enhancer and D/V-boundary enhancer in *cis*, activates *lacZ*-reporter gene in

a

```

TATGCCCTC CGGAGACCGG GGGCCCAAAA ATAGCAACTG CAATTGAGCG GCAGGAGATA CCAAAAACTT GCATAGGCTT GCATTTGCGT TGACAACATT 100
CCAAACTCGA TTCGAGAGAA AAATATCCAA CAACTGGAGA GGAGTTTTGA GGATGCGGAC GAGGACGAGG TCGGAGGATG TGGATGTGGA TGTGCTTGTG 200
GGGATGTGTA ATGTTTGTGC TTGGCTGCCG TCGGATTCG ACACTTTGG CCGGCACGTT GGCAGGTGTG CCATGCATGC TGATGACGAT GAGAATGAGG 300
ATGAGGATGA GGA GCGGAT GATGATGGTG CTGGTCCGGG TGGATACTGA ATACTGGATA CCGGATGCCA TGCCGCTGC CTTTTTTTTC CCGTACCAGA 400
AGCCAGAAAT CGTCATCATC CCATTGCCAT TCACTCACTC GCTCAGCTGA GGCCTGGGAA TTCCCATTA TGTGCAACA AGCAGACTGC CAAAGATATT 500
TCCTCTGCAG CTCCCTCAGT TAGCATTTCA CTTTCAGCCA GCGGTTTCAA AAGCGAAACG CGCAGTACCT GTCCCACTT TCACCTTTTG GCTTAATGAA 600
GAGAGCGTGG CGATTTATGA CCCGATAACG TTCGATCGCC AGCGTTGACG CATAGTCCGG TCGTCACAG AGAAACTATC CTAATCCTCT ATGAAGATTG 700
TAAATCAAGT ATGGAGTATT CAAACTAAAA TCGACTCAAC AGCAGTTTTA TCTTGCTTAA ACTATAGTGA GTTTCAATTA GAAACGACAA AGTACAATCG 800
CTAGAG 806

```



the complete pattern of the endogenous *vg* expression pattern (compare with Fig. 5a).

METHODS. See ref. 26 for antibody staining methods.

wing discs (compare Figs 3d and e; enhancer and *vg* expression in the notum is not Su(H)-dependent). This indicates that this single Su(H) binding site is essential for the *vg* D/V enhancer function, making *vg* the first N–Su(H)-regulated target gene to be identified that is outside of the neural–epidermal lateral inhibition pathway.

A second enhancer for the growing wing blade

The *vg* protein is expressed and required in all cells that will give rise to the wing blade and hinge which form a 'pouch' in the third instar imaginal disc. As the Su(H)-regulated D/V boundary enhancer controls only part of the native *vg* pattern, there must be other *cis*-regulatory sequences controlling *vg* expression in wing-pouch cells other than the D/V boundary cells. We searched throughout the *vg* locus for such enhancer elements and found that a 806-bp DNA segment (Fig. 4a) in the fourth intron of the *vg* gene is sufficient to drive *lacZ* reporter gene expression in all of the pouch cells except those along the D/V boundary (Fig. 4b). The pattern of reporter gene expression is bisected by non-expressing D/V boundary cells and becomes lower along the A/P boundary in the late third instar, giving a quadrant-like pattern. Unlike the D/V boundary enhancer which is active in the second instar³⁸, the 'quadrant' enhancer is not active until the early third instar (Fig. 4c–f). Interestingly, the quadrant enhancer is initially activated in a small patch of cells near the intersection of the A/P and D/V compartment boundaries (Fig. 4d–f) after the D/V boundary is established and when the imaginal disc and wing primordium start to grow rapidly. This suggests that signals from both the A/P and D/V boundaries may be involved in quadrant-enhancer activation.

To determine whether the combination of the quadrant enhancer and the D/V boundary enhancer is sufficient to account for the native *vg* pattern, we joined the two elements in *cis* and placed them upstream of *lacZ*. This reporter construct gives *lacZ* expression in the complete spatial and temporal pattern of the native *vg* gene in the wing disc (Fig. 4g).

Quadrant enhancer requires D/V organizer

The complementary patterns of the D/V boundary and quadrant enhancers suggest that the two elements are either responding to

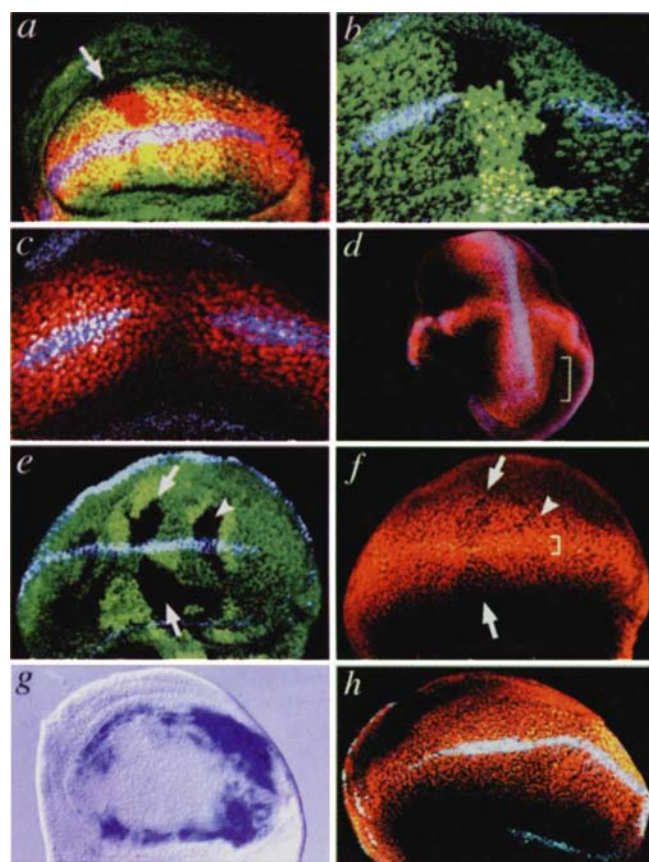
different cues or are regulated differently by the same cues. To examine the regulation of the quadrant enhancer and of the native *vg* gene (and hence wing-cell proliferation) in cells away from the D/V boundary, we analysed gene expression in wing discs bearing patches of mutant cells, or in which transgenes were expressed ectopically, concentrating on signalling pathways that influence wing formation in the larval stages, for example, the *Notch*^{27,28,39}, *dpp*¹⁶, and *wingless* pathways^{40,41}. It should be noted that as *vg* is absolutely required for wing-cell proliferation (Fig. 1), any pathway that is required for *vg* activation is expected to yield no, or extremely small and short-lived, null mutant clones within the presumptive wing.

To test the effect of the N-receptor signalling pathway, we generated *Su(H)*[−] clones in the wing discs and examined *vg* protein and quadrant-*lacZ* reporter gene expression. Throughout the developing wing blade, away from the D/V boundary, *Su(H)*[−] clones had no effect on *vg* gene expression, indicating that quadrant enhancer activation is *Su(H)*-independent (Fig. 5a). This is consistent with previous reports that *N* function is not required for formation of this part of the wing during the third instar⁴².

There is a striking non-autonomous effect of *Su(H)*[−] clones at the D/V boundary on *vg* expression in cells surrounding the clone. If *Su(H)*[−] clones cross (data not shown) or touch the D/V boundary from the ventral side (Fig. 5b, c), the level of *vg* expression drops as the distance of individual cells from wild-type boundary cells increases. This suggests that a signal from cells at the D/V boundary, the production of which depends on the N–*Su(H)* pathway, affects *vg* expression in the quadrants away from the D/V boundary. The importance of D/V boundary cells for *vg* activation in the presumptive wing blade is also supported by three other observations. First, generation of ectopic boundary-type cells by the ectopic expression of an activated *Notch* receptor (*N*^{int}) induces wing outgrowth^{28,43} and *vg* expression in both *N*^{int}-expressing and adjacent wild-type cells (Fig. 5d). Second, *ap* mutants that lack a D/V boundary lack all *vestigial* expression, not just that controlled by the D/V boundary enhancer⁴⁰. Third, in the *vg*^{83b27} mutant which lacks the D/V boundary enhancer³⁸ and fails to form D/V boundary cells, no *vg* expression is induced

FIG. 5 The quadrant enhancer is activated by a signal from the D/V boundary and the *dpp* pathway. The expression of the Myc epitope tag is shown in green (b and e); *vg* protein in red (a, c, d, f and h); *wg* protein in blue (b and c); β -galactosidase driven by the *vg* D/V-boundary enhancer in magenta (a, d). a, The *Su(H)*⁻ null clones in non-D/V-boundary cells have no effect on *vg* expression (arrow). Co-localization of *vg* protein and Myc epitope tag is shown in yellow. b, A *Su(H)*⁻ clone (blank area) touching the D/V boundary ventrally abolishes the *wg* expression (shown in blue) at both sides of the D/V boundary. This clone can grow because it still expresses *vg* because of derepression of the quadrant enhancer (see c). Repression of the quadrant enhancer at the D/V boundary is *Su(H)*-dependent and probably mediated by *E(spl)m8* (see a), c, *vg* (red) and *wg* (blue) expression in the disc shown in b. Note that the pouch cells around the D/V-boundary cells, which do not express *wg* in the *Su(H)*⁻ clone show a lower level of *vg* protein than normal. In contrast, the *vg* level is higher in cells that are closest to the *wg*-expressing D/V boundary cells which are not affected by the *Su(H)*⁻ clone. d, Ectopic expression of a constitutively active N (intra-N) in the *patched* pattern along the A/P boundary not only autonomously activates the *vg* D/V-boundary enhancer (magenta) but also non-autonomously activates *vg* in the nearby cells (red, see bracket). e, Myc epitope tag staining showing *tkv* null clones harbouring one copy of a *tkv* transgene (arrows and arrowhead). *wg* expression (blue) marks the D/V boundary and wing-pouch area. The twin spot cells have two copies of the wild-type *tkv* gene and one copy of the *tkv* transgene. The rest of the cells have one copy of each wild-type *tkv* and *tkv* transgene. f, *vg* expression in the same disc shown in e. *vg* expression is reduced in the *tkv* null clones with one copy of *tkv* transgene (arrows). Note that *vg* expression in the D/V-boundary cells in the *tkv* clone crossing the D/V boundary (arrowhead) is not affected (bracketed area). g, *UAS-lacZ* expression pattern driven by *GAL4*⁶²⁶ driver. h, *vg* expression in *UAS-dpp* discs driven by *GAL4*⁶²⁶. The *vg*-expressing area and level of *vg* protein greatly increased in the posterior side of the pouch (right) where the *dpp* protein level is higher (compare with 'g').

METHODS. The *tkv* clones for panels e and f were generated by heat-shock induction of FLP recombinase in *w¹¹¹⁸, hsFLP1; tkv⁵, P[ry⁺, hs-neo, FRT]40A/P[w⁺, hs- π M]36F; P[w⁺, *tkv*^{Q1A8}]/+* flies.



(J. Williams and S.B.C., unpublished data). Importantly, ectopic expression of *vg* is incapable of inducing *vg* expression in adjacent cells (not shown). These results, and the temporal order of D/V boundary and quadrant enhancer activation suggest that D/V-boundary cells must produce a signal (or signals) that organize the rest of the wing.

One candidate D/V-organizing signal is the *wingless* (*wg*) protein which is activated at the D/V boundary by the *N-Su(H)* pathway²⁶⁻²⁹. However, we have not found any evidence to suggest that *wg* regulates *vg* in the developing wing blade. Unlike ectopic expression of the activated *N* receptor (Fig. 5d), ectopic expression of *wg* does not alter *vg* levels (data not shown) or cause wing overgrowth (A. Martinez-Arias, personal communication). Furthermore, large *wg*⁻ or *dishevelled* mutant clones that run along either side of the D/V boundary have little effect on wing growth²⁸⁻⁴⁴. The striking difference in ectopic *N^{trunc}* versus ectopic *wg* expression suggests that there is another *N-Su(H)*-activated factor involved in wing outgrowth. The discovery of the quadrant enhancer should help to elucidate the identity of the *Su(H)*-dependent wing pouch-organising signal(s).

Activation of quadrant enhancer by *dpp*

The wings of adult viable *dpp* mutant flies are defective in growth as well as in patterning¹⁶, and *dpp* has been proposed to produce the organizing signal for A/P patterning and growth in *Drosophila* imaginal discs^{13,16-19}. Ectopic expression of *dpp* is sufficient to induce overgrowth of imaginal discs and adult wings, and to reorganize A/P patterning elements^{13,17-19}. As *vg* mutant flies also have wing-disc growth defects, and ectopic *vg* expression causes outgrowths of wing tissue, we have examined the potential regulatory relationship between the *dpp* signalling pathway and *vg*-gene activation.

The *thick veins* (*tkv*) gene encodes a type-I Dpp receptor⁴⁵⁻⁴⁷. Flies in which *tkv* function is reduced have smaller wings and wing discs, and these mutants exhibit reduced *vg*-protein levels and cell

number in the pouch, away from the D/V boundary (data not shown). To examine the effect of loss of *tkv* function, we examined *tkv*⁻ clones. Importantly, *tkv*⁻ null clones in the wing pouch do not survive beyond the two-cell stage⁴⁸, whereas those outside the wing pouch grow well (A. Penton and M. Hoffmann, personal communication). To circumvent the difficulties presented by tiny *tkv*⁻ clones, we induced *tkv*⁻ clones in flies bearing a *tkv* transgene⁴⁹ that rescues cell viability. In *tkv*⁻ clones of this background, cells in the developing wing blade exhibit lower levels of *vg* protein expression (Fig. 5e, f, arrows) but those at the D/V boundary do not (Fig. 5e, f, arrowhead). Whereas reduction in *dpp* signalling reduces *vg* protein levels, ectopic *dpp* expression induces *vg* expression (Fig. 5g, h). Taken together, these data suggest that *vg* is downstream of the *dpp* pathway and may mediate its growth-related function.

vestigial regulates wing growth and identity

The phenotype of *vg* mutant flies, the *vg*-expression pattern, the selective requirement for *vg* in proliferating wing cells, the trans-determination of other adult structures into wing by targeted expression of *vg*, and the direct regulation of the *vestigial* gene by compartmental sources that are required for wing formation, all point to *vestigial* as the central regulator of wing-cell behaviour and identity. We have examined the role of one other effector gene, *scalloped* (*sd*)⁴⁰, in wing growth and differentiation, but it seems to play a subordinate role to that of *vestigial*. We have found that targeted expression of *sd* does not induce proliferation in imaginal discs, wing outgrowths in adult tissues or *vg* expression (A.S. K. Irvine, and S.B.C., unpublished observations). Targeted expression of *vg* does, however, induce *sd* expression, so it is possible that *sd* is regulated by *vg* and may control a subset of the cellular events downstream of *vg*.

The function of *vestigial* in wing development seem analogous to the recently described function of the *eyeless* gene in eye development⁵⁰. Both genes are necessary for the formation of an

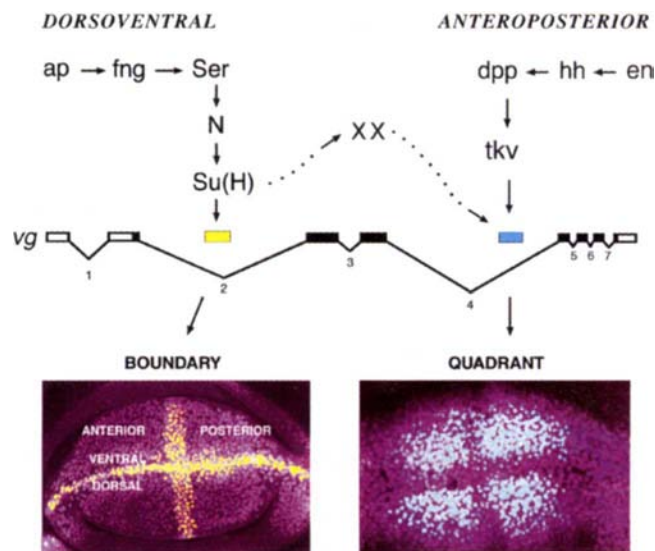


FIG. 6 The *vg* gene contains two *cis*-regulatory elements that control complementary sectors of the *vg* expression pattern in the developing wing. The 'boundary' enhancer in the second intron (left, yellow block) is activated at the D/V boundary directly through the N-Su(H) pathway. The 'quadrant' enhancer in the fourth intron (right, blue block) is activated in the developing wing blade by means of signal(s) emanating from the D/V boundary (XX, centre) and *dpp* protein from the A/P boundary (top right).

entire structure and sufficient to induce characteristic cell types in ectopic locations. However, neither the ancestry nor the function of *vestigial* is as clear as for *eyeless/Pax-6*. The *eyeless/Pax-6* gene is ancient, appearing very early in metazoan evolution and subsequently controlling the development of light-sensitive structures in diverse phyla. It would be interesting to examine the phylogenetic distribution and function of the *vestigial* gene in primitive winged and wingless insects, as well as other arthropods and animals.

Signal integration at the *vestigial* gene

We have shown that the *vg* gene contains two separate *cis*-regulatory elements that control complementary sectors of the native *vg* pattern and are regulated by different signalling pathways. The *vg* D/V-boundary enhancer in the second intron is activated early in wing development by the D/V signalling pathway, which is regulated by the *apterous* dorsal selector gene through the *Ser* ligand, *Notch* receptor, and *Su(H)* transcription factor (Fig. 6, left). The enhancer in the fourth intron is expressed later in the quadrants of the developing wing blade and responds to signals generated by both A/P (*dpp*) and D/V-boundary cells (signal XX; Fig. 6, right). From the architecture of the regulatory pathways that lead to *vg* activation, it can now be appreciated how loss of *ap*, *fng*, *Ser*, *N*, *Su(H)*, *en*, *hh*, or *dpp*, results in failure of wings to form by virtue of loss of *vg* activation. One prediction of this regulatory model is that targeted expression of *vg* should restore wing outgrowth in mutants lacking these gene activities, as we confirmed.

It has been understood for some time that in animal appendages, the elaboration of the proximodistal axis, orthogonal to the primary A/P and D/V axes of the embryo, requires information initially provided by these two axes. Various models (such as the boundary⁵¹ or polar coordinate⁵² models) have focused on formal rules for the integration of this axial information in the formation and growth of the circularly organized appendage fields. However, the molecular mechanisms underlying signal integration and growth have not been deciphered for any appendage. The regulatory elements within the *vg* gene suggest that signals from boundary sources in the developing wing first establish the D/V organizer and then regulate wing growth and identity by means of signal integration from both the D/V and A/P sources. In both general function and respective position, the D/V organizer of the wing disc is analogous to the AER of the vertebrate limb bud and the posterior compartment of the wing disc is analogous to the ZPA which produces the A/P-organizing signal(s). It will be interesting to learn how genes that are expressed in the mesenchyme of the progress zone, and depend upon signals from both the AER and ZPA, integrate these signals during the growth and patterning of the vertebrate limb. □

Received 9 April; accepted 10 June 1996.

1. Tabin, C. *Cell* **80**, 671–674 (1995).
2. Riddle, R. D. et al. *Cell* **75**, 1401–1416 (1993).
3. Cohn, M. J. et al. *Cell* **80**, 739–746 (1995).
4. Parr, B. A. & McMahon, A. P. *Nature* **374**, 350–353 (1995).
5. Yang, Y. & Niswander, L. *Cell* **80**, 939–947 (1995).
6. Riddle, R. D. *Cell* **83**, 631–640 (1995).
7. Laufer, E. et al. *Cell* **79**, 993–1003 (1994).
8. Niswander, L. et al. *Nature* **371**, 609–612 (1994).
9. Crossley, P. H. et al. *Cell* **84**, 127–136 (1996).
10. Blair, S. S. *BioEssays* **17**, 299–309 (1995).
11. Lee, J. J. et al. *Cell* **71**, 33–50 (1992).
12. Tabata, T., Eaton, S. & Kornberg, T. B. *Genes Dev.* **6**, 2635–2645 (1992).
13. Basler, D. & Struhl, G. *Nature* **368**, 208–214 (1994).
14. Tabata, T. & Kornberg, T. G. *Cell* **76**, 89–102 (1994).
15. Sanicola, M. et al. *Genetics* **139**, 745–756 (1995).
16. Posakony, L., Raftery, L. & Gelbart, W. *Mech. Dev.* **33**, 69–82 (1991).
17. Capdevilla, J. & Guerrero, I. *EMBO J.* **6**, 715–729 (1994).
18. Ingham, P. & Feitz, M. J. *Curr. Biol.* **5**, 432–440 (1995).
19. Zecca, M., Basler, K. & Struhl, G. *Development* **121**, 2265–2278 (1995).
20. Diaz-Benjumea, F. & Cohen, S. M. *Cell* **75**, 741–752 (1993).
21. Cohen, B. et al. *Genes Dev.* **6**, 715–729 (1992).
22. Blair, S. et al. *Development* **120**, 1805–1815 (1994).
23. Irvine, K. & Wieschaus, E. *Cell* **79**, 595–606 (1994).
24. Sprecher, S. et al. *Development* **120**, 535–544 (1994).
25. Fleming, R. et al. *Genes Dev.* **4**, 2188–2201 (1990).
26. Kim, J., Irvine, K. D. & Carroll, S. B. *Cell* **82**, 795–802 (1995).
27. Couso, J. P., Knust, E. & Martinez Arias, A. *Curr. Biol.* **5**, 1437–1448 (1995).
28. Diaz-Benjumea, F. J. & Cohen, S. *Development* **121**, 4215–4225 (1995).
29. de Celis, J. F., Garcia-Bellido, A. & Bray, S. J. *Development* **122**, 359–369 (1996).
30. Simpson, P., Lawrence, P. A. & Maschat, F. *Dev. Biol.* **84**, 206–211 (1981).
31. Brand, A. H. & Perrimon, N. *Development* **118**, 401–415 (1993).
32. di Pasquale, A. *Ro. Ist. Lomb. Sci. Lett. III* **85**, 241–248 (1952).
33. Ouweneel, W. J. *Genetica* **41**, 1–20 (1970).

34. Fortini, M. E. & Artavanis-Tsakonas, S. *Cell* **79**, 273–282 (1994).
35. Lecourtis, M. & Schweisguth, F. *Genes Dev.* **9**, 2598–2608 (1995).
36. Bailey, A. M. & Posakony, J. W. *Genes Dev.* **9**, 2609–2622 (1995).
37. Xu, T. & Rubin, G. M. *Development* **117**, 1223–1237 (1993).
38. Williams, J. et al. *Nature* **368**, 299–305 (1994).
39. Couso, J. P. & Martinez Arias, A. *Cell* **79**, 259–272 (1994).
40. Williams, J. A. et al. *Development* **117**, 571–584 (1993).
41. Couso, J. P., Bate, M. & Martinez-Arias, A. *Science* **259**, 484–489 (1994).
42. de Celis, J. F. & Garcia-Bellido, A. *Mech. Dev.* **46**, 109–122 (1994).
43. Doherty, D. et al. *Genes Dev.* **10**, 421–434 (1996).
44. Axelrod, J. D. et al. *Science* **271**, 1826–1832 (1996).
45. Nellen, D. et al. *Cell* **78**, 225–237 (1994).
46. Penton, A. et al. *Cell* **78**, 239–250 (1994).
47. Brummel, T. J. et al. *Cell* **78**, 251–261 (1994).
48. Nellen, D. et al. *Cell* **85**, 357–368 (1996).
49. Morimura, S. et al. *Dev. Biol.* (in the press).
50. Halder, G., Callaerts, P. & Gehring, W. J. *Science* **267**, 1788–1792 (1995).
51. Meinhardt, H. *Semin. dev. Biol.* **2**, 129–138 (1991).
52. French, V., Bryant, P. J. & Bryant, S. V. *Science* **193**, 969–981 (1976).

ACKNOWLEDGEMENTS. J.K. and A.S. contributed equally to this work. We thank A. Bailey and J. Posakony for a generous gift of the *Su(H)* protein; D. Brower, M. Hoffmann, G. Halder and U. Hinz for GAL4 driver strains; A. Penton, Y. Chen and M. Hoffmann for UAS-*dpp* and *tkv* receptor stocks; L. Maves for suggestions on tissue mounting; K. Irvine for the UAS-*scalloped* line; A. Martinez-Arias, E. Rulifson and S. Blair for valuable information on *wg*, *sgg*, and *dsh*; R. Nusse for *wg* antibody; D. Doherty for UAS-*N^{tr}* flies; F. Schweisguth for *Su(H)* mutant flies; J. Williams for preliminary work; J. Gates for technical assistance; S. Paddock for assistance with confocal-microscopy; G. Pangniban, D. Lewis, M. Hoffmann, A. Laughon, and A. Penton for comments on original versions of the manuscripts; and J. Wilson for preparing the manuscript. A.S. is a NIH predoctoral trainee in genetics. This work was supported by the Howard Hughes Medical Institute.

CORRESPONDENCE and requests for materials should be addressed to S.B.C. (e-mail: sbcarroll@facstaff.wisc.edu).

Heated gaseous streamers and star formation in the Orion molecular cloud

Jennifer J. Wiseman*† & Paul T. P. Ho*

* Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, Massachusetts 02138, USA

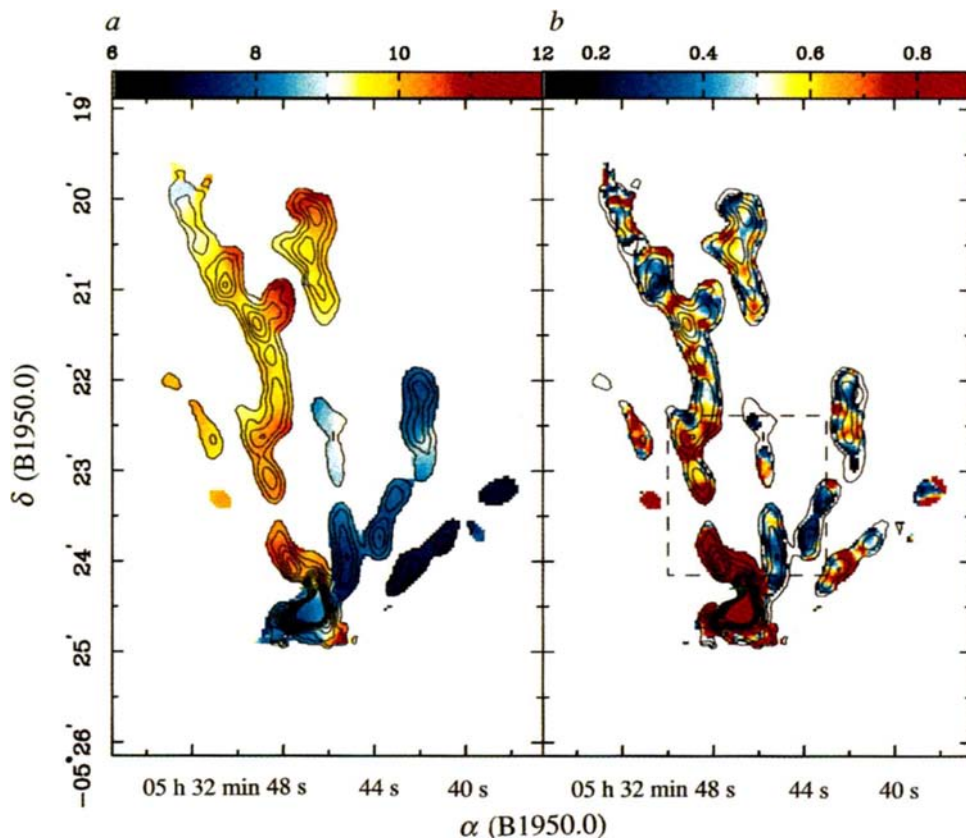
† National Radio Astronomy Observatory, 520 Edgemont Road, Charlottesville, Virginia 22903, USA

THE Orion molecular cloud, which is obscured by the dust and ionized gas of the Orion nebula, is the nearest example of a giant molecular cloud. Massive stars are actively forming deep in the core of this cloud as a result of large-scale cloud instabilities, fragmentation and gravitational collapse. These young stars will inject a considerable amount of energy back into the surrounding environment through stellar winds and radiation¹, and they are thus expected to exert a major influence on the evolution of the cloud. Here we present a mosaic of ten high-resolution radio maps of the region of the cloud known as OMC-1; the maps were constructed from observations of two ammonia emission lines, which trace the densest regions of the gas while mitigating the obscuring effects of the dust. We find dense filaments of molecular gas with complex motions fanning out more than 0.5 parsec from the central core of the cloud. These filaments appear as long, bead-like chains, consisting of dense clumps of gas that may be the sites of future star formation. The outer sheaths of clumps and the edges of filaments may be heated as a direct result of radiation and outflows from young stars embedded in the central core.

Previous high-angular-resolution ($<10''$) maps of molecular line emission in OMC-1 have been largely limited to the $1'$ (0.15 pc) region around the embedded sources in the central core, Orion-KL. Larger-scale maps, spanning arc minutes or degrees, have revealed the structure of the cloud to be filamentary with embedded dense cores²⁻⁷, but detailed examination of the structure, kinematics and temperature on such a large scale have not previously been possible because of instrumental and mapping limitations. Our high-resolution maps were obtained by observing the NH_3 (J, K) = (1,1) and (2,2) inversion transitions with the Very Large Array (VLA), operated by the National Radio Astronomy Observatory, in its compact 'D' configuration, with newly improved receivers. Twenty adjacent fields across the OMC-1 ridge were synthesized, individually deconvolved to remove the defects from aperture synthesis, and then corrected for primary-beam attenuation. For each of 20 velocity channels spaced by 0.3 km s^{-1} , all fields were linearly combined to produce a large $1.2 \text{ pc} \times 0.6 \text{ pc}$ mosaic with 0.02 pc ($8''$) spatial resolution. The mosaics of the channel maps for each NH_3 transition were then used in combination to obtain integrated intensity, spectra and kinematic information. The ammonia spectra are very useful for deriving physical parameters in the molecular cloud. The transitions observed here produce spectra with a strong central component and satellite components at both higher and lower frequencies. The intensity ratios between the main and satellite components provide a simple measure of the optical depth. This, together with the line intensity, linewidth and the fractional abundance of NH_3 in the J, K state, allows determination of the ammonia column density. Densities and masses can then be calculated by adopting estimates for the ammonia abundance and cloud geometry⁸.

The remarkable structure of the OMC-1 core can be seen in Fig. 1. The contours represent the intensity of the NH_3 (1,1) emission, integrated across the linewidth of the transition, with $8''$

FIG. 1 The OMC-1 Ridge. The maps were made from ten adjacent VLA fields in the NH_3 (1,1) and (2,2) inversion transition lines. The fields were separated by $1'$ and observed with an $8''$ synthesized beamwidth. The Orion-KL high-mass star-forming core is the southern core in the image; the position of the outflow source IRc2 is marked with a cross. *a*, The contours show the intensity of the central (1,1) line integrated from 6.5 to 12.5 km s^{-1} . The shades portray the velocity field by displaying the first-moment integral over velocity; the velocity resolution is 0.3 km s^{-1} . Note the high-velocity gradients in the northern cores and along some filament edges. Note also the global bimodal velocity field, with 10 km s^{-1} material to the east, 6 – 8 km s^{-1} material to the west, and an overlapping region near Orion-KL. *b*, The contours show, as before, the integrated intensity of the main (1,1) line emission. The shades display the (2,2)/(1,1) intensity ratio, which is directly related to the temperature of the region. The broken lines outline the area displayed in Fig. 2. Note the strong heating near Orion-KL, and the edge heating and inter-clump heating elsewhere.



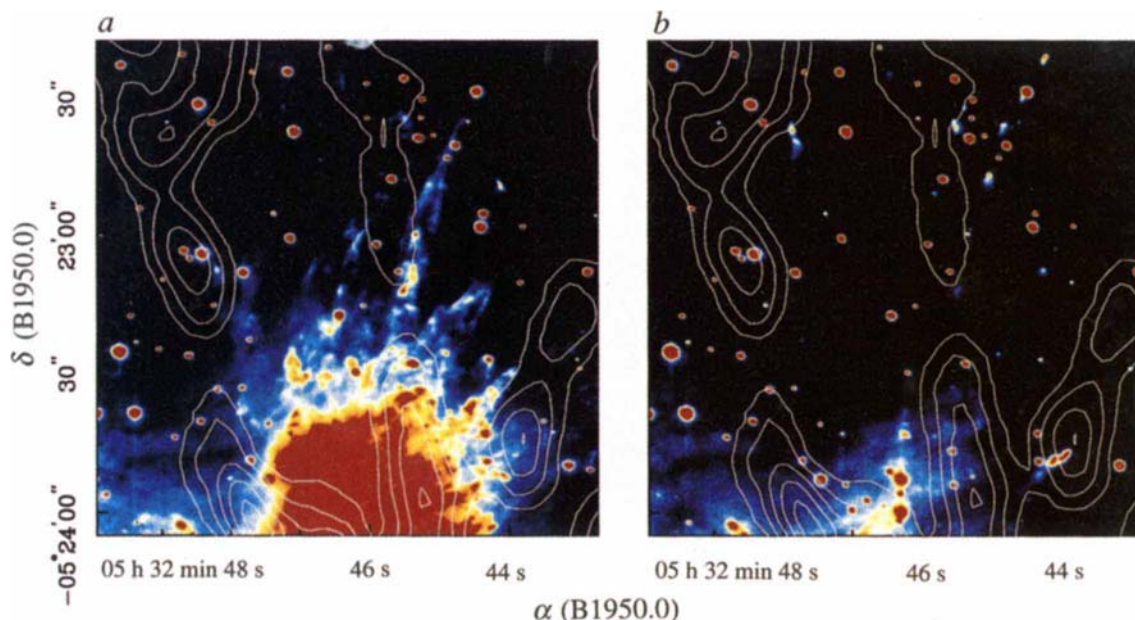


FIG. 2 Overlay of the NH_3 emission with the H_2 (a) and $[\text{Fe II}]$ (b) shock emission maps of Allen and Burton²². Circular spots are field stars. The $[\text{Fe II}]$ emission knots appear at the positions of the 'tips' of several elongated H_2 fingers. Note how the shocked emission fingers fill the lower lobe within the NH_3 filament region in the same orientation. Note also the proximity of the

$[\text{Fe II}]$ knot at $\alpha = 05^{\text{h}} 32^{\text{m}} 48^{\text{s}}$, $\delta = -05^{\circ} 22' 35''$ to the NH_3 filament edge region, which seems to be heated (see Fig. 1b). Outflows and ejecta from embedded sources or from unstable swept-up shells might be energetically impacting the surrounding molecular environment.

resolution. From the Orion-KL dense core region at $\alpha = 05^{\text{h}} 32^{\text{m}} 47^{\text{s}}$, $\delta = -05^{\circ} 24' 24''$ (the position of the infrared cluster and outflow source IRC2 (ref. 9)), the emission fans out into a number of long filaments that extend to >0.5 pc from the core. The filaments seem to consist of clumps at roughly regular spacings, like beads on a necklace. The optical depth of ammonia ranges from ~ 0.5 between the clumps to 1 or 2 at the clump peaks.

It is important to know whether this ammonia emission pattern is tracing the distribution of mass, the bulk of which is in the form of invisible molecular hydrogen. Ammonia is typically excited in the denser cores of molecular clouds, and it is used along with other molecules like CS to map the distribution of gas near regions of star formation^{8,10}. But molecular emission can also be enhanced or decreased by inhomogeneities in either the fractional abundance of the molecule or the chemical excitation conditions across the region^{11,12}. In some cases, emission from the ammonia transitions used here seems to correspond very well with the structures seen in other trace molecules¹⁰, whereas in other regions, particularly very hot peaks towards exciting stars, the ammonia emission seems to be depleted¹¹. In OMC-1, previous lower-resolution maps of NH_3 (ref. 13) and HC_3N (refs 7,14) show a similar pattern of clumpy filaments. As a more precise check, we have also mapped the central $2.5' \times 1'$ central region at $15''$ resolution in the $^{12}\text{C}^{18}\text{O } J=1-0$ line, a reliable tracer of mass distribution, using the Haystack 37-m single-dish radio telescope¹⁵. In the central core, the peak of the $^{12}\text{C}^{18}\text{O}$ emission is offset by $\sim 30''$ from the NH_3 peak. On moving out from the core, the $^{12}\text{C}^{18}\text{O}$ corresponds strikingly well to the NH_3 , with both tracers rising and falling in intensity together along the pattern of clumps and filaments. Thus the extended morphology of OMC-1 seems to be correctly characterized as elongated, 'beaded' streamers.

This structure suggests large-scale gravitational instabilities in the detected gas distribution. Theoretical calculations have shown that elongated cylinders of compressible material are gravitationally unstable, fragmenting into condensations with spacing determined by the mass distribution, angular momentum and magnetic field effects¹⁶⁻¹⁸. Fragmentation has been seen before in the larger Orion complex on a $20'$ scale in C^{18}O maps³; we now see that this hierarchical pattern of clumping of material within filaments

continues down to angular scales of $\sim 30''$. This is an important size scale for low-mass star formation, because the masses of these individual clumps range between 1 and $4M_{\odot}$. In particular, some clumps that are seen in the maps appear to be self-gravitating and may be the sites of future stars.

The ammonia mosaic also reveals a startling new glimpse of motions and heating in the filaments. In Fig. 1a the shades represent the velocity field of the region, with the eastern dense filamentary material at a velocity with respect to the local standard of rest of $\sim 10 \text{ km s}^{-1}$, the northwestern filaments at $\sim 8 \text{ km s}^{-1}$ and a western filament at 6 km s^{-1} . The former two velocity components overlap markedly in the core region near the position of the outflow source IRC2. An examination of the spectrum at this position shows a clear blending of these two velocity features. This suggests that this system of filaments may be in direct physical contact here, possibly even triggering instabilities and star formation. The origin of the differential motions of the filaments is unclear¹⁹. As the larger Orion complex became unstable and collapsed into filaments, the conservation of angular momentum might have produced the general velocity gradient between the filament components. Another possibility is the direct impact of outflows from the central embedded star cluster, with obvious heating effects as discussed below. A related possibility is that filaments might line the walls of an expanding cone surrounding a collimated outflow, creating different projected radial filament velocities. Some of the constituent clumps along the filaments have strong velocity gradients. The northernmost clumps, especially, show fast gradients that indicate a binding mass of approximately $3M_{\odot}$ each, if they are gravitationally bound and rotating. They might be newly collapsed cores that will become new star formation sites if they can lose their angular momentum.

A striking temperature pattern is apparent, too. The shades of Fig. 1b represent the temperature of the gas as deduced from the ratio of the intensities of the (2,2) and (1,1) transitions, with the more shaded (red) regions being warmer. Along the filaments the temperature generally falls to $\sim 20 \text{ K}$ within emission peaks and rises to $>50 \text{ K}$ between them. This argues for external heating of the core sheaths, from either foreground stars or the embedded central cluster at Orion-KL. The evidence supports the postulate

of Hartquist and Dyson²⁰ that interface regions between dense cores and surrounding winds are likely sites for turbulence and prolonged heating. The filament edges also show strong heating effects near the southern section of the main ridge (eastern) filament, facing IRC2 to the south, as well as in the partial 'ring' surrounding IRC2 (ref. 7). These heated features might delineate the approximate boundary of the northern lobe of outflowing molecular gas from the IRC2 region; shocks and shearing from the outflowing material could heat and shred the filament edges. Radiation travelling through the evacuated outflow region from the embedded sources in the core might also reach and heat these exposed edges²⁰. For sources of $10^4 L_{\odot}$, heating of the dust to temperatures of 50 K can be easily achieved out to a radius of 0.5 pc under ordinary conditions²¹. Provided the gas is well coupled to the dust, the gas will also be heated.

A comparison of Fig. 1a and 1b shows that some edge regions that are heated also show some velocity gradients, as though they were being 'pushed'. Taken together, these maps of large-scale filamentary structure, internal motion and widespread heating in OMC-1 portray a model cycle of vibrant interaction between a molecular cloud core and the stars that form within. An unstable cloud region has collapsed into interacting filaments that are fragmenting further into bead-like chains of clumps, some of which are collapsing and 'spinning up'. Winds and radiation from young stars are heating the sheaths of these clumps. Outflows from massive young stellar objects forming within the central core are shredding, shocking and heating the filament edges, inducing instability and returning still more energy to the environment.

The nature of the impact of outflows on the surrounding cloud can be clarified with observations of tracers of shocked gas. There is a remarkable correlation between the molecular filaments with their heated edges and the distribution of shocked 'fingers' of H₂ 2.122 μ m and [Fe II] 1.644 μ m emission as reported by Allen and Burton²², including Herbig-Haro (H-H) knots with high-velocity proper motions^{23,24}. Figure 2 shows the overlay of the shocked H₂ (Fig. 2a) and [Fe II] (Fig. 2b) emission with the NH₃ filaments. The H₂ fingers fill the region within the cone defined by the outermost (east-west) NH₃ emission. It is apparent that the outflow from IRC2 follows the same orientation as the molecular filaments; the shock emission here might be the direct result of the outflow's impacting and shocking the filament surfaces. The stronger [Fe II] emission is seen at the finger 'tips' as H-H objects, possibly the result of the deceleration of gaseous clumps ejected from the central star. Alternatively, Stone *et al.*²⁵ propose that these 'bullets' are actually the result of Rayleigh-Taylor instability and fragmentation of a dense shell swept up by an outflow that is time-variable. If the filament region is situated on the near edge of the molecular cloud⁷, then the steep density gradient could also contribute to the effect. The [Fe II] knot at $\alpha = 05^{\text{h}} 32^{\text{m}} 48^{\text{s}}$, $\delta = -05^{\circ} 22' 35''$ has a particularly striking projected spatial correspondence with a region of edge heating along the eastern NH₃ filament (see Fig. 1b). We could be observing the effects of direct impacts of a fragmented shell swept up by an outflow from a distant embedded source 0.3 pc away.

Future studies of this region will employ higher excitation lines, which will probe the heated interface regions between the outflow and the molecular filaments. Higher-resolution observations of both structure and kinematics are also needed, because the continued stability of filamentary structure is predicted to depend on the relationship between the velocity field and density enhancements along the length of the individual filaments²⁶. The new millimetre and sub-millimetre wavelength telescope arrays under development should provide the capability for high-resolution probes of material within the densest regions and should aid in the detection of any further fragmentation²⁷. $\square$

Received 5 February; accepted 21 May 1996.

- Shu, F. H., Adams, F. C. & Lizano, S. A. *Rev. Astr. Astrophys.* **25**, 23–81 (1987).
- Bailly, J., Langer, W., Stark, A. & Wilson, R. *Astrophys. J.* **312**, L45–L49 (1987).
- Dutrey, A. *et al. Astr. Astrophys.* **247**, L9–L12 (1991).
- Maddalena, R., Morris, M., Moscovitz, J. & Thaddeus, P. *Astrophys. J.* **303**, 375–391 (1986).

- Batrla, W., Wilson, T. L., Bastien, P. & Ruf, K. *Astr. Astrophys.* **128**, 279–290 (1983).
- Tatematsu, K. *et al. Astrophys. J.* **404**, 643–662 (1993).
- Rodríguez-Franco, A., Martín-Pintado, J., Gómez-González, J. & Planesas, P. *Astr. Astrophys.* **264**, 592–609 (1992).
- Ho, P. T. P. & Townes, C. H. A. *Rev. Astr. Astrophys.* **21**, 239–270 (1983).
- Dougados, C., Léna, P., Ridgway, S. T., Christou, J. C. & Probst, R. G. *Astrophys. J.* **406**, 112–121 (1993).
- Zhou, S., Wu, Y., Evans, N., Fuller, G. & Myers, P. *Astrophys. J.* **346**, 168–179 (1989).
- Davis, C. J. & Dent, W. R. F. *Mon. Not. R. astr. Soc.* **261**, 371–378 (1993).
- Ungerechts, H. *et al. Lect. Not. Phys.* no. 459 (eds Winnewisser, G. and Petz, G. C.) 258–264 (Springer, Heidelberg, 1995).
- Murata, Y. *et al. Astrophys. J.* **359**, 125–130 (1990).
- Martín-Pintado, J., Rodríguez-Franco, A. & Bachiller, R. *Astrophys. J.* **357**, L49–L52 (1990).
- Wiseman, J. & Ho, P. T. P. in *Clouds, Cores, and Low Mass Stars* (eds Clemens, D. P. and Barvainis, R.) 396–400 (Conf. Ser. 65, Astr. Soc. Pacif., San Francisco, 1994).
- Stodólkiewicz, J. S. *Acta astr.* **13**(1), 30–64 (1963).
- Larson, R. B. *Mon. Not. R. astr. Soc.* **214**, 379–398 (1985).
- Hanawa, T. *et al. Astrophys. J.* **404**, L83–L86 (1993).
- Wornack, M., Ziurys, L. M. & Sage, L. J. *Astrophys. J.* **406**, L29–L33 (1993).
- Hartquist, T. W. & Dyson, J. E. Q. *Jl. R. astr. Soc.* **34**, 57–71 (1993).
- Ho, P. T. P., Martin, R. N. & Barrett, A. H. *Astrophys. J.* **246**, 761–787 (1981).
- Allen, D. A. & Burton, M. G. *Nature* **363**, 54–56 (1993).
- Axon, D. J. & Taylor, K. *Mon. Not. R. astr. Soc.* **207**, 241–261 (1984).
- Taylor, K. N. R., Storey, J. W. V., Sandell, G., Williams, P. M. & Zealey, W. J. *Nature* **311**, 236–237 (1984).
- Stone, J. M., Xu, J. & Mundy, L. G. *Nature* **377**, 315–317 (1995).
- Gehman, C. S., Adams, F. C., Fatuzzo, M. & Watkins, R. *Astrophys. J.* **457**, 718–730 (1996).
- Ishiguro, M. & Welch, W. J. (eds) *Astronomy with Millimeter and Submillimeter Wave interferometry Sect. 4*, 384–431 (Conf. Ser. 59, Astr. Soc. Pacif., San Francisco, 1994).

ACKNOWLEDGEMENTS. We thank M. Burton for the use of infrared images of Orion, and J. Tedds for assistance with these images. NRAO is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc.

CORRESPONDENCE to be addressed to J.J.W. (e-mail: jwiseman@nrao.edu).

A strongly magnetic neutron star in a nearly face-on binary system

Pascal Daumerie*, Vassiliki Kalogera†, Frederick K. Lamb*† & Dimitrios Psaltis*†

* Physics Department, University of Illinois at Urbana-Champaign, 1110 West Green Street, Urbana, Illinois 61801, USA

† Astronomy Department, University of Illinois at Urbana-Champaign, 1002 West Green Street, Urbana, Illinois 61801, USA

THE first example of a new type of transient X-ray source recently appeared in the direction of the Galactic Centre¹. During the peak of its outburst, the new source, GRO J1744 – 28, was very bright in X-rays² and produced both 2.1-Hz periodic pulsations³ and intense bursts lasting tens of seconds^{1,4}. Before the discovery of this source, it was thought that X-ray stars could not display these different types of activity simultaneously. Here we discuss the nature of the source, which seems to be a strongly magnetic neutron star accreting matter from a low-mass companion star in a low-inclination orbit. The dipole component of its magnetic field is $\lesssim 10^{11}$ G. When the source was at its brightest, its X-ray luminosity between bursts was close to the Eddington critical luminosity⁵, at which the outward force of the escaping radiation balances the inward force of gravity. The X-ray bursts probably occur when matter that has accumulated in the inner part of the accretion disk briefly overcomes the forces that oppose its inflow, and the gas falls onto the neutron star.

In order to understand the pulsar's spin rate and how it produces X-ray bursts and periodic pulsations, we need to know (1) the strength of the pulsar's magnetic field, (2) the pulsar's luminosity, and (3) the nature of the binary system. The pulsar's magnetic field funnels the inflowing matter towards the magnetic poles, producing beams of X-rays that sweep across our line of sight as the pulsar spins. The strength of the pulsar's magnetic field, together with its spin frequency and accretion rate, determine whether centrifugal force is important in the outer magnetosphere. The pulsar's luminosity determines whether the outward force of the escaping radiation can successfully oppose the inflow

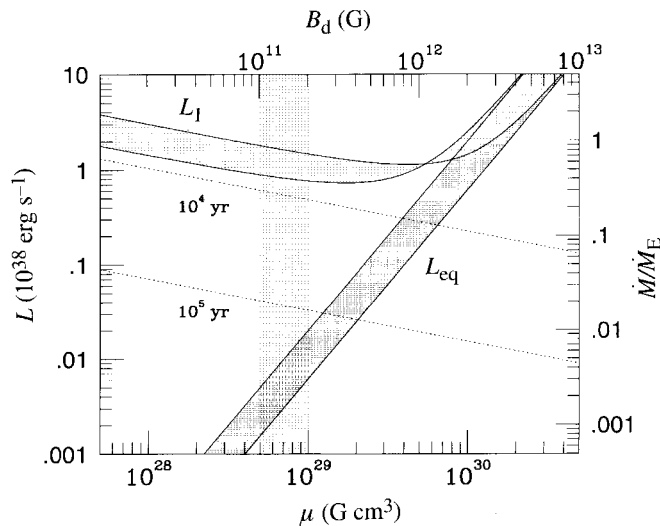


FIG. 1 Relations between the pulsar's luminosity (L) and magnetic dipole moment (μ). The shaded curved band labelled L_1 shows the mean accretion luminosity that would produce the spin-up rate observed⁷ on 1 January 1996, as a function of μ . This curve demonstrates that L_1 was greater than $8 \times 10^{37} \text{ erg s}^{-1}$. The darker diagonal band labelled L_{eq} shows the 'equilibrium' accretion luminosity at which the standard theory of disk accretion predicts that the accretion torque vanishes, as a function of μ . The vertical scale on the right shows the corresponding mass accretion rates in units of $\dot{M}_E$, the accretion rate that would produce a neutron star luminosity equal to the Eddington critical luminosity; the horizontal scale at the top shows the dipole magnetic field at the pulsar's pole (B_d). The lightly shaded vertical band shows the upper limit on μ derived in the text. The diagonal dotted lines show where in the diagram the spin-evolution timescale is 10^4 and 10^5 yr. All results are for matter of cosmic composition accreting onto a neutron star of radius 10 km with an inertial moment $I = 0.3 MR^2$. The luminosity bands were determined by solving numerically the magnetohydrodynamic equations for a slim disk interacting with a strongly magnetic neutron star (P.D., F.K.L. and P. Ghosh, manuscript in preparation). Such a computation is needed to treat accurately a rapidly rotating star with a luminosity near the Eddington critical luminosity, but the results are similar to those obtained earlier using a simpler approach⁶, except that the fastness (ω_s ; see equation (1)) at which the torque vanishes is 0.7 rather than 0.3. The widths of the bands indicate the uncertainty in the results introduced by the uncertainty in the mass of the pulsar and in the strength of the magnetic coupling between the pulsar and the disk; where $L \gtrsim 2 \times 10^{38} \text{ erg s}^{-1}$, radiation forces introduce additional uncertainty.

of matter. The nature of the binary system governs the average mass-transfer rate and determines the evolution of the pulsar's spin and magnetic field.

We can derive an upper bound on the dipole magnetic moment of the pulsar and estimate its luminosity L from its spin-up rate, because infall of matter produces a torque on the pulsar that depends on its dipole magnetic moment and accretion rate. According to the standard theory⁶ of disk accretion by magnetic stars, the rate of change of the spin frequency ν_s of a star of mass M , radius R , and moment of inertia I , with a dipole magnetic field B_d at the magnetic pole, is

$$\dot{\nu}_s = 6n(\omega_s)(M/M_\odot)^{-10/7} B_{12}^{2/7} R_6^{-2/7} (MR^2/I) L_{38}^{6/7} \text{ pHz s}^{-1} \quad (1)$$

Here $B_{12} \equiv B_d/(10^{12} \text{ G})$, $R_6 \equiv R/(10^6 \text{ cm})$, $L_{38} \equiv L/(10^{38} \text{ erg s}^{-1})$, and $n(\omega_s)$ is a dimensionless function of the fastness parameter $\omega_s \equiv \nu_s/\nu_K(\bar{\omega}_0)$, which measures the stellar spin frequency in terms of the keplerian frequency $\nu_K(\bar{\omega}_0)$ at the cylindrical radius $\bar{\omega}_0$ of the inner edge of the keplerian flow. If $\dot{\nu}_s$ is set equal to 8.6 pHz s^{-1} , its value⁷ on 1 January 1996, equation (1) establishes a relation between the dipole component of the pulsar's magnetic field and L_1 , its mean luminosity at this epoch (see Fig. 1).

Figure 1 shows that L_1 is $\gtrsim 10^{38} \text{ erg s}^{-1}$ and hence that in January the pulsar's luminosity between bursts was near the Eddington critical luminosity, which is $2 \times 10^{38} \text{ erg s}^{-1}$ for matter of cosmic composition accreting onto a star of mass 1.4 solar

masses ($1.4 M_\odot$; ref. 5). This luminosity is consistent with the 2–200 keV X-ray flux measured by the Rossi Explorer satellite⁸ if the radiation is beamed by a factor ~ 4 and the source is $\sim 7 \text{ kpc}$ away. This distance is consistent with preliminary measurements^{8,9} of the column density of neutral hydrogen between us and the pulsar and the properties of the proposed soft X-ray and infrared counterpart of the X-ray star¹⁰.

The pulsar's brightness declined steadily after about 14 January and its spin-up rate decreased, following closely (M. Finger, personal communication) the $\dot{\nu}_s \propto L^{6/7}$ dependence predicted by equation (1) for a gas-pressure-dominated inner disk. Spin-down had not begun by 6 May, at which time the pulsar was at least 170 times fainter than on 1 January (M. J. Stark, personal communication), implying that the equilibrium luminosity L_{eq} at which the accretion torque vanishes is more than 170 times smaller than L_1 and that the dipole component of the pulsar's magnetic field is $\lesssim 10^{11} \text{ G}$ (Fig. 1). If spin-down began shortly after 6 May, the pulsar's dipole magnetic field is $\sim 10^{11} \text{ G}$ and its peak luminosity between bursts at the height of its outburst was $\sim (1-3) \times 10^{38} \text{ erg s}^{-1}$. The dipole component of the pulsar's magnetic field cannot be much less than $\sim 10^{11} \text{ G}$ or the inner edge of the disk would have become radiation-pressure-dominated at the height of the outburst, causing $\dot{\nu}_s$ to have a different dependence on L .

The periodic pulsations at the pulsar's spin frequency vary in amplitude from $\sim 10\%$ at energies of a few keV (ref. 9) to $\sim 40\%$ at energies above 30 keV (refs 3, 4), indicating that the fundamental cyclotron energy is $\sim 10 \text{ keV}$ and hence that the surface magnetic field is $\sim 10^{12} \text{ G}$. Also, the pulsar's X-ray spectrum⁸ is very similar to the spectra¹¹ of other accreting neutron stars that have magnetic

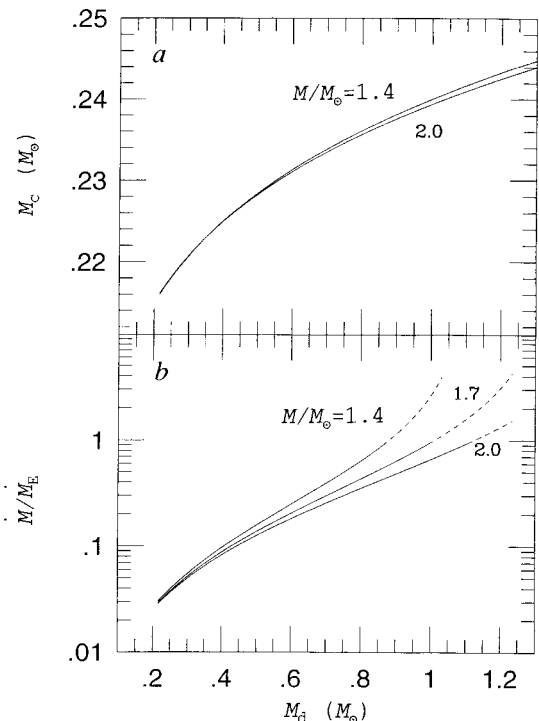


FIG. 2 a, Critical core mass (M_c) of a giant-branch donor star that just fills its Roche lobe in a binary with $P_{\text{orb}} = 11.8 \text{ d}$, as a function of the total mass of the donor (M_d), for two different neutron-star gravitational masses M . b, Rate of mass transfer driven by nuclear evolution of a low-mass giant companion as a function of the total mass of the donor, in units of $\dot{M}_E$, the accretion rate that would produce a neutron-star luminosity equal to the Eddington critical luminosity, for matter of cosmic composition accreting onto a 10^6 cm -radius neutron star. The curves show results for three different neutron-star gravitational masses and have been calculated assuming conservative mass transfer; deviations from conservative mass transfer are expected when $\dot{M} > \dot{M}_E$ (dashed portions of the curves).

fields $\sim 10^{12}$ G. Thus, the pulsar's surface magnetic field may be much stronger than its dipole component.

This analysis indicates that GRO J1744–28 has a strong magnetic field and a luminosity between bursts that, at the peak of its outburst, was at or above the Eddington critical luminosity. The pulsar SMC X-1 is known to be a strongly magnetic neutron star accreting near the critical rate and has been seen¹² to produce an X-ray burst similar to the many bursts produced by GRO J1744–28, but it is in a binary system with a high-mass companion. It is very unlikely that GRO J1744–28 has a high-mass companion because its mass function is very small⁷, so the system would have to have an improbably small inclination ($i \leq 1.5^\circ$ for a donor mass $M_d \geq 10 M_\odot$, which has an *a priori* probability $\leq 3 \times 10^{-4}$). Also, there is no reason for a young, high-mass star to lie in the direction of the Galactic Centre, where most stars have low masses. Thus, GRO J1744–28 almost certainly has a low-mass companion.

The Doppler shift of its pulse frequency shows that GRO J1744–28 is in a binary system with an orbital period of 11.8 days (ref. 7). A low-mass star in such a wide system can fill its gravitational potential well (Roche lobe) and transfer matter to the neutron star only if it had a mass of at least $1 M_\odot$, when it was on the main sequence and has now become a giant¹³. Stellar evolution calculations¹⁴ show that the radius of such an evolved, low-mass star depends primarily on the mass of its degenerate core. Hence we can use the core mass–radius relation for such stars, together with Newton's laws, to compute the critical core mass M_c of a donor star that just fills its Roche lobe in this system, as a function of its total mass M_d and the mass M of the neutron star. Figure 2a shows that M_c depends only weakly on M_d and hardly at all on M . The mass of the companion cannot be less than $0.22 M_\odot$, which is the mass of an evolved star with the critical core mass and a negligible envelope mass. On the other hand, mass transfer would be dynamically unstable and would swamp the neutron star if the donor were more massive than $\sim 1.5 M_\odot$, taking into account non-conservative mass transfer¹³. Therefore the mass of the donor must be between 0.22 and $1.5 M_\odot$. These bounds on the companion mass and the measured X-ray mass function imply that we are viewing the system nearly face-on ($i < 23^\circ$ for $M < 2.0 M_\odot$), contrary to what has been assumed previously^{7,15}. The rate of mass transfer from the donor star to the pulsar, averaged over the binary evolution time, depends primarily on the mass of the donor, as shown in Fig. 2b.

The current spin frequency GRO J1744–28 reflects the accretion torque, and hence the accretion rate, averaged over its spin evolution time, which is $\sim 10^6$ yr (Fig. 1). Stated differently, the mean luminosity of GRO J1744–28 over the past 10^6 yr must have been about equal to the luminosity $L_{eq} \approx 10^{36}$ ergs⁻¹ that would produce an equilibrium spin frequency equal to the current spin frequency (Fig. 1). If the system is nearing the end of the mass transfer phase, the companion has transferred ~ 0.5 – $0.8 M_\odot$ to the pulsar and now has a mass ~ 0.2 – $0.5 M_\odot$, and the mass accretion rate averaged over the past 10^6 yr was about an order of magnitude smaller than the mass transfer rate averaged over the binary evolution time (Fig. 2b). Moreover, the pulsar will never become a millisecond pulsar.

Our analysis indicates that GRO J1744–28 is a strongly magnetic neutron star. The large X-ray bursts are probably produced when matter that has accumulated in the inner part of the accretion disk briefly overcomes the forces that oppose its inflow and surges onto the neutron star (Fig. 3). After such an inward surge of matter, the inflow rate (and hence the X-ray flux) is likely to be depressed until the accretion flow has recovered. This picture of the cause of the bursts is supported by four observations: (1) the similarity of the X-ray spectrum during the bursts to the X-ray spectrum between bursts^{8,9}, which indicates that the energy sources and emission processes during and between bursts are the same; (2) the wide variety of burst durations^{1,16}, which is expected if they are produced by unstable accretion; (3) the change in the pulse phase that occurs during and

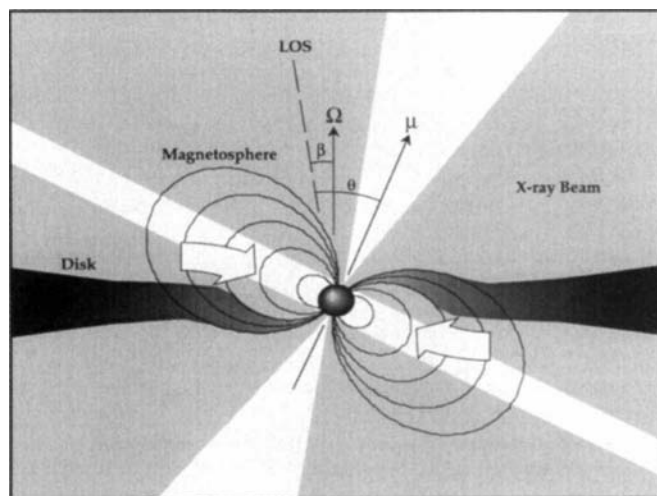


FIG. 3 Cartoon (not to scale) showing a side view of the model for the bursting pulsar described in the text. The accretion torque quickly aligns the pulsar's spin axis Ω with the axis of the accretion disk, which is approximately normal to the orbital plane of the binary system, so the spin axis is now normal to the orbital plane. As shown in the text, we see the system almost face-on, so our line of sight (LOS) is close to (but not exactly along) the pulsar's spin axis. The pulse waveform of GRO J1744–28 is very nearly sinusoidal over the entire energy range from 2 to 200 keV and there appears to be only one peak per rotation⁹. This waveform is unique among known pulsars, but is to be expected when the angle β between our line of sight and the pulsar spin axis is small. Then the variation δ with rotation phase ϕ of the angle θ between our line of sight and the magnetic axis (μ) is also small, and the X-ray flux we see will vary as $F(\phi) \approx F(0) + \delta(\phi)F'(0) \approx F(0) + \delta_{\max} \sin(\phi)F'(0)$, that is, approximately sinusoidally. X-ray bursts occur when matter surges from the accretion disk onto the neutron star (large arrows).

after a burst^{17,18}, which indicates a shift in the flow pattern onto the pulsar; and (4) the reduction in the X-ray flux that occurs for $\sim 1,500$ s following each burst^{8,9,18}.

If this picture is correct, then first, at the height of its outburst the luminosity of the pulsar between bursts was $\sim (1-3) \times 10^{38}$ ergs⁻¹ (which implies that the beaming factor is ~ 4 if it is ~ 7 kpc away). Second, the pulsar's dipole magnetic field is $\lesssim 10^{11}$ G. Third, the pulsar will begin to spin down if its mean luminosity falls below $L_{eq} \approx 10^{38} B_{12}^2$ ergs⁻¹, which is ~ 270 times lower than its luminosity at the height of its outburst for $B_d \approx 10^{11}$ G. Fourth, quasi-periodic X-ray oscillations (QPOs) at the beat-frequency between the pulsar's spin frequency and the keplerian frequency at the inner edge of the accretion disk^{19,20} should have a frequency ~ 30 Hz at the height of the outburst and ~ 1 Hz when the pulsar begins to spin down. The 20, 40, and 60 Hz QPOs that have been observed²¹ cannot be the beat-frequency QPO because their frequencies remained unchanged as the luminosity of the pulsar fell by a factor ~ 12 (W. Zhang, personal communication). The properties of the beat-frequency QPO should change substantially during each burst. Fifth, if the magnetic field at the surface of GRO J1744–28 is $\sim 10^{12}$ G, much stronger than the dipole component there, cyclotron scattering lines may be found in the 10–20 keV energy range. Because cyclotron lines have not yet been detected, the variation of $\dot{v}$, with luminosity and the frequency of the beat-frequency QPO may be the only ways to determine accurately the magnetic field of this unique object.

Note added in proof: We have recently become aware of a paper by Finger *et al.*²² that was submitted after ours and which provides additional information on this source. $\square$

Received 25 January; accepted 7 May 1996.

1. Kouveliotou, C. *et al.* *Nature* **379**, 799–801 (1996).

2. Paciesas, W. S., Harmon, B. A., Fishman, G. J., Zhang, S. N. & Robinson, C. R. *IAU Circ.* No. 6284 (1995).

3. Finger, M. H. et al. *IAU Circ. No.* 6285 (1996).
4. Kouveliotou, C. et al. *IAU Circ. No.* 6286 (1996).
5. Lamb, F. K. & Miller, M. C. *Astrophys. J.* **439**, 828–845 (1995).
6. Ghosh, P. & Lamb, F. K. *Astrophys. J.* **234**, 296–316 (1979).
7. Finger, M. H., Wilson, R. B. & van Paradijs, J. *IAU Circ. No.* 6286 (1996).
8. Swank, J. *IAU Circ. No.* 6291 (1996).
9. Dotani, T. et al. *IAU Circ. No.* 6337 (1996).
10. Kouveliotou, C. et al. *IAU Circ. No.* 6369 (1996).
11. Mészáros, P. *High Energy Radiation from Magnetized Neutron Stars* (Univ. Chicago Press, 1992).
12. Angelini, L., White, N. E. & Stella, L. *Astrophys. J.* **371**, 332–341 (1991).
13. Kalogera, V. & Webbink, R. F. *Astrophys. J.* **458**, 301–311 (1996).
14. Webbink, R. F., Rappaport, S. & Savonije, G. J. *Astrophys. J.* **270**, 678–693 (1983).
15. Lewin, W. H. G., Rutledge, R. E., Kommers, J. M., van Paradijs, J. & Kouveliotou, C. *Astrophys. J.* **462**, L39–L42 (1996).
16. Giles, B. & Strohmayer, T. *IAU Circ. No.* 6338 (1996).
17. Jung, G. V. et al. *IAU Circ. No.* 6321 (1996).
18. Stark, M. J., Baykal, A., Strohmayer, T. & Swank, J. H. *IAU Circ. No.* 6334 (1996).
19. Lamb, F. K., Shibazaki, N., Alpar, M. A. & Shaham, J. *Nature* **317**, 681–687 (1985).
20. Shibazaki, N. & Lamb, F. K. *Astrophys. J.* **318**, 767–785 (1987).
21. Zhang, W. et al. *IAU Circ. No.* 6300 (1996).
22. Finger, M. H. et al. *Nature* **381**, 291–293 (1996).

ACKNOWLEDGEMENTS. We thank R. Rothschild and J. Swank for providing XTE results in advance of publication, and D. Lamb, C. Miller, J. Swank and R. Webbink for helpful comments. This research was supported in part by NASA and the US NSF.

A crystal structure determined with 0.02 Å accuracy by electron microscopy

Thomas E. Weirich*, Reiner Ramlau*, Arndt Simon*, Sven Hovmöller† & Xiaodong Zou‡

* Max-Planck-Institut für Festkörperforschung, Postfach 80 06 65, D-70506 Stuttgart, Germany

† Structural Chemistry, Stockholm University, S-106 91 Stockholm, Sweden

‡ Mineralogy and Petrology, Institute of Geology, Lund University, Sölvegatan 13, S-223 62 Lund, Sweden

ELECTRON crystallography has two important advantages over X-ray crystallography for the determination of atomic positions in crystal structures: crystallographic phases can be determined directly from images¹, and extremely small samples can be analysed. On the other hand, the strong interaction between electrons and matter gives rise to dynamical effects² which complicate quantitative analysis of the experimental data. This has led to a pessimistic view of the possibility of direct crystal structure determination by electron crystallography^{3,4}, especially for compounds containing heavy elements. The atomic structures of a few inorganic crystals, mainly oxides, have been determined from electron microscopy images^{5–7}, but in no case was the atomic structure refined. Here we report the complete determination of an unknown structure, for the compound $\text{Ti}_{11}\text{Se}_4$, by electron crystallography. These results show that crystals that are too small for single-crystal X-ray diffraction and difficult to solve by powder diffraction may nevertheless be amenable to accurate structure determination by electron crystallography.

Metal-rich compounds frequently contain clusters with metal–metal (M–M) bonded octahedral M_6 cores. Depending on the relative metal content, these octahedral units may condense, that is, share corners, edges or faces. The condensed M_6 clusters can be regarded as slightly distorted fragments of the body-centred cubic (b.c.c.) metal itself. Structures containing condensed metal clusters exhibit interesting structural and physical properties. The metal–metal bonds may extend over regions from zero to three dimensions. The conductivity ranges from metallic to semiconducting, and also includes superconductivity.

The general principle of condensed clusters⁸, seen in the long-known structure of Ti_8S_3 (ref. 9), indicated that this compound was probably the first of many related structures. Recently, metal-rich compounds in the Ti–Se system have been prepared. Amongst

these, crystals large enough for structure determination by single-crystal X-ray diffraction have been obtained of Ti_9Se_2 (ref. 10), Ti_8Se_3 (ref. 11) and Ti_7Se (ref. 12). Their structures are based on Ti_6 octahedra which are deformed as in the b.c.c. close-packed arrangement. These octahedra share edges and corners, to form frameworks of condensed clusters.

In the course of investigations of the Ti–Se system by electron microscopy, another new compound was found, which proved to be $\text{Ti}_{11}\text{Se}_4$. An attempt to solve the structure from X-ray powder diffraction data failed because the sample was inhomogeneous and the structure very complex. Crystals large enough for single-crystal X-ray diffraction could not be obtained. Thus, electron crystallography was the method of choice to carry out the full structure analysis. Within X-ray crystallography, a crystal structure is considered correctly and accurately determined only after a successful refinement, with several times more experimental data points (diffraction amplitudes) than unknowns (atomic coordinates and temperature factors), leading to good geometry and a low R -value. We wanted to see if a crystal structure could be refined against electron-diffraction amplitudes to give a result comparable with X-ray diffraction techniques. The sample size needed for high-resolution electron microscopy (HREM) and crystallographic image processing is about 100 unit cells, and for electron diffraction about 10,000 unit cells; the latter is several million times smaller than the size needed for X-ray crystallography, even using synchrotron radiation.

The lattice parameters of $\text{Ti}_{11}\text{Se}_4$ were determined by selected-area electron diffraction (SAED) and later refined from X-ray powder data to $a = 25.516(11)$ Å, $b = 3.4481(14)$ Å, $c = 19.201(6)$ Å, $\beta = 117.84(3)^\circ$. SAED patterns showed a C -centred monoclinic unit cell. The intensity distributions of $(h0l)$ and $(h2l)$ were essentially identical, from which we concluded that all atoms must be located on mirror planes perpendicular to the short b -axis, as is the case with Ti_8Se_3 (ref. 11) and the isostructural Ti_8S_3 (ref. 9). Thus only the space groups Cm and $C2/m$ are possible. Chemical microanalysis by energy-dispersive X-ray spectroscopy

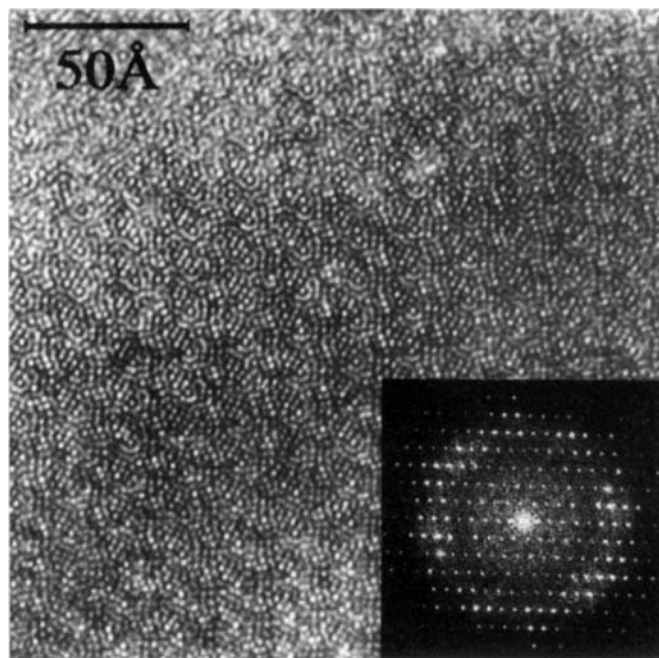


FIG. 1 HREM image of $\text{Ti}_{11}\text{Se}_4$ taken along the short b -axis with a 300-kV Philips CM30/ST electron microscope. The image was digitized using a video-rate CCD (charge-coupled device) camera and analysed by the image processing system CRISP¹⁰. An area of 512×512 pixels (corresponding to $190 \text{ Å} \times 190 \text{ Å}$) at the edge of the crystal was selected and its Fourier transform (FT) calculated (inset). The lattice of diffraction spots was refined, and amplitudes and phases extracted from this calculated FT.

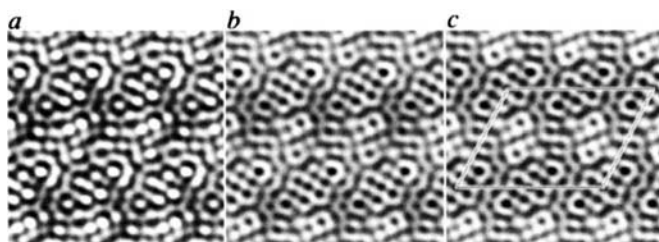


FIG. 2 Projection maps of $\text{Ti}_{11}\text{Se}_4$ at various stages of image processing. *a*, Only lattice averaging. Here white dots are atoms (reversed contrast). *b*, After compensating for the CTF, the contrast is correct and atoms are seen as black peaks. *c*, After imposing the crystallographic symmetry in projection, $p2$. All the 90 atoms in the unit cell (marked) were localized from this map. These maps were calculated using all 113 diffraction spots out to 1.75-Å resolution, by inverse Fourier transformation of the amplitudes and phases extracted⁹ from the FT. Correct contrast was restored by dividing each pixel in the FT by the CTF¹⁷. No pixels were amplified more than four times to avoid magnifying uncertain data near the cross-over. Division by a negative number of the CTF implies a reversal of phase. The CTF was calculated for defocus -930 Å , spherical aberration constant $C_s = 1.15\text{ mm}$, an estimated beam convergence of 0.069° and defocus spread of 70 Å .

(EDXS)¹³ in the electron microscope gave a Ti content of 72(1) atomic per cent.

HREM images are directly interpretable in terms of structure only if the experimental conditions are optimal. Two criteria must be met: the crystals must be very thin to minimize multiple scattering (weak phase objects), and the images must be taken at the optimal defocus, the so-called Scherzer defocus. Then, all information out to the point resolution of the microscope is transmitted with correct contrast. In this case Scherzer defocus was at -560 Å and the point resolution of the microscope was 1.9 Å . Electron optical effects, which prevent the direct interpretation of the images, are described by the contrast transfer function (CTF). This function has the effects of changing the

image (crystallographic) phases, equivalent to reversing the contrast, and attenuating amplitudes.

An image of a thin crystal (Fig. 1) was processed. An area near the edge was selected and its Fourier transform (FT) calculated (inset). From the radius of the first cross-over of the CTF, seen as a dark ring in the FT, the defocus was determined to be -930 Å . The information outside the dark ring at 3.9 Å d -spacing has reversed contrast. After compensating for the effects of the CTF, atoms appear with their correct contrast (Fig. 2). The crystal symmetry for the $(h0l)$ -projection was determined from analysis of the phases in the experimental image. The plane groups of the two possible space groups Cm and $C2/m$ are $p1$ and $p2$, respectively. All phases must be 0° or 180° in a projection with $p2$ symmetry when the unit cell origin is placed on a twofold axis, whereas there are no phase restrictions in $p1$. An origin refinement search gave a phase residual of only 18° for $p2$. Thus the crystal symmetry was $C2/m$.

The phases of all $(h0l)$ reflections extracted from the FT were set to the closest of 0° or 180° and a new projection map was calculated (Fig. 2c), showing 90 peaks per unit cell. Although the peak heights did not differentiate between the two atom types in the structure, the 23 unique peaks could be assigned to 17 Ti and 6 Se atoms following the well established principles of cluster condensation⁸. All the Ti atoms are part of condensed metal octahedra, while the Se atoms are surrounded by Ti atoms forming trigonal prisms. Atomic coordinates were measured directly from the projection map. The coordinates in the third dimension (along y) were assigned 0 or $\frac{1}{2}$ from geometrical considerations. The empirical formula derived in this way, $\text{Ti}_{11}\text{Se}_4$, has 73.3 atomic percent Ti, close to the 72 ± 1 found by EDXS.

The 44 positional parameters (x - and z -coordinates for all 23 unique atoms except one Ti which is fixed in the special position $(0,0,0)$) and 23 isotropic temperature factors were refined against electron diffraction intensities, quantified from SAED patterns taken along the b -axis of a relatively thin crystal (Fig. 3). The SHELXL-93¹⁴ least-squares refinement program with atomic scattering factors for electrons¹⁵ was used. The crystallographic R -value was 48.4% before and 14.7% after refinement, for all 408 unique non-zero reflections to 0.75-Å resolution. No correction was applied for multiple scattering or curvature of the Ewald sphere. The final R -value of only 14.7% is remarkably good for

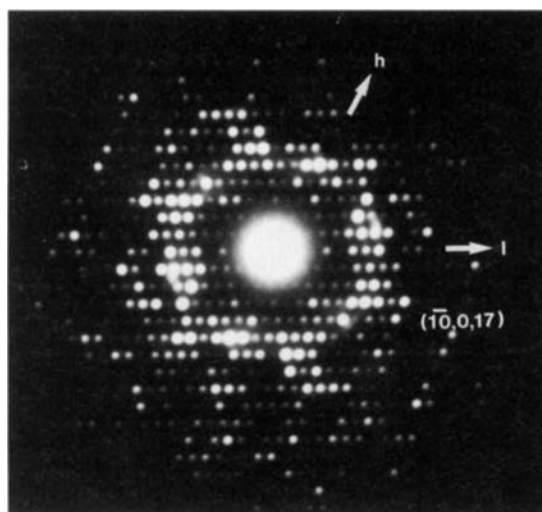


FIG. 3 Electron-diffraction pattern of $\text{Ti}_{11}\text{Se}_4$ along the b -axis. A series of five SAED patterns with exposure times ranging from 0.5 to 7.5 s were recorded on photographic film and quantified. The SAED pattern shown here was exposed 7.5 s. The weakest reflections were only measurable on the longer exposures, while the strongest reflections were overexposed on these. The films were digitized on a light-box, using a video-rate CCD camera. The diffraction intensities were extracted using the ELD computer program¹⁸. 408 unique reflections to 0.75-Å resolution were collected from this projection. The extracted intensities were used as $F_{\text{obs}}^2(h0l)$ for the refinement. A strong reflection at 1.13-Å resolution $(10,0,17)$ is marked.

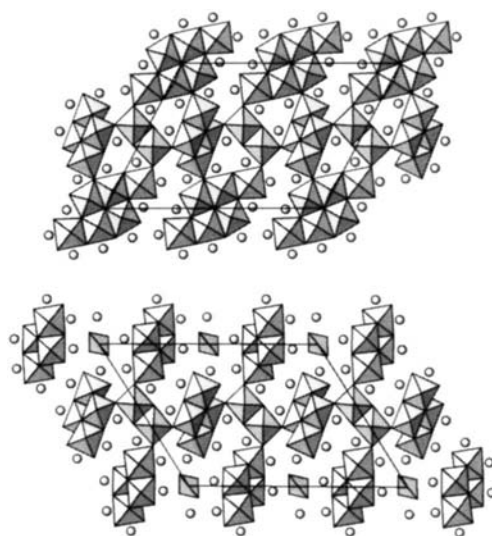


FIG. 4 Structure model of (top) $\text{Ti}_{11}\text{Se}_4$ obtained by electron crystallography and (bottom) the related structure Ti_6Se_3 solved by single-crystal X-ray diffraction¹¹. Ti_6 octahedral clusters are shown as part-shaded four-sided polygons, and Se atoms are shown as circles. The central layer of alternating strings of two and four Ti octahedra parallel to the ab -plane is identical in both structures.

electron-diffraction data, and argues strongly that dynamical effects play at most a minor part in influencing the diffraction amplitudes. This is even more significant when we consider that this crystal contains no light elements and, as judged from the transparency of the crystals, were not thinner than normal for specimens used for HREM and SAED investigations. The samples were prepared by standard techniques; they were crushed in an agate mortar, and the thinnest crystals were selected for imaging and diffraction analysis. Electron-diffraction patterns from other crystals were all similar to the series used here. We did not determine the thickness of the crystals, but estimate them to be between 50 and 150 Å thick, which are quite normal values. (For more detailed crystallographic data, see Supplementary Information.)

The atoms shifted on average by 0.2 Å in the refinement. The standard deviations of refined atomic positions calculated by SHELXL-93 ranged from 0.015 to 0.021 Å. The average accuracy in atomic positions obtained here by electron crystallography (0.02 Å) is close to what may be expected from refinement against X-ray diffraction data. The temperature factors were about twice as high for the Se atoms as for the Ti atoms.

The structure of $\text{Ti}_{11}\text{Se}_4$ has many features in common with the other structures in the Ti–Se system. In particular, $\text{Ti}_{11}\text{Se}_4$ and Ti_8Se_3 are closely related, as seen from Fig. 4, which shows the underlying common construction set. Note that every second layer consists of alternating strings of two and four condensed octahedra chains in both structures. The distortions of the octahedra in the equivalent motifs are remarkably similar. This is a very strong argument in favour of the reliability of our refinement.

Electron crystallography opens a window into the vast world of

microcrystalline compounds, natural as well as synthetic, including minerals, alloys and other new materials. With slight modifications, electron crystallography might be further extended to provide atomic positions in ordered structures which are not crystalline in the traditional sense. These include defects, interfaces and quasi-crystals. □

Received 14 March; accepted 30 May 1996.

- DeRosier, D. J. & Klug, A. *Nature* **217**, 130–134 (1968).
- Cowley, J. M. *Diffraction Physics* 2nd edn (North-Holland, Amsterdam, 1984).
- Atkins, P. W. *Physical Chemistry* 5th edn (Oxford Univ. Press, 1994).
- Gjønnes, J. & Steeds, J. W. in *International Tables for Crystallography* Vol. C, (ed. Wilson, A. J. C.) 363–365 (Kluwer Academic, Dordrecht, 1992).
- Hovmöller, S., Sjögren, A., Farrants, G., Sundberg, M. & Marinder, B.-O. *Nature* **311**, 238–241 (1984).
- Downing, K. H., Hu, M., Wenk, H.-R. & O'Keefe, M. A. *Nature* **348**, 525–528 (1990).
- Hovmöller, S. *Trans. Am. Cryst. Ass.* **28**, 95–103 (1992).
- Simon, A. *Angew. Chem. int. Edn. engl.* **20**, 1–22 (1981).
- Owens, J. P. & Franzen, H. F. *Acta Crystallogr.* **B30**, 427–430 (1974).
- Weirich, T. E., Simon, A. & Pöttgen, R. *Z. Anorg. Allg. Chem.* **622**, 630–634 (1996).
- Weirich, T. E., Pöttgen, R. & Simon, A. *Z. Kristallogr.* (in the press).
- Weirich, T. E., Pöttgen, R. & Simon, A. *Z. Kristallogr.* (in the press).
- Goldstein, J. J. in *Introduction to Analytical Electron Microscopy* (eds Hren, J. J., Goldstein, J. I. & Joy, D. C.) 83–120 (Plenum, New York, 1979).
- Sheldrick, G. M. *Acta Crystallogr.* **A46**, 467–473 (1990); SHELXL-93, *Program for Crystal Structure Refinement* (Göttingen, 1993).
- Doyle, P. A. & Turner, P. S. *Acta crystallogr.* **A24**, 390–397 (1968).
- Hovmöller, S. *Ultramicroscopy* **41**, 121–136 (1992).
- Zou, X. D., Sundberg, M., Larine, M. & Hovmöller, S. *Ultramicroscopy* **62**, 103–121 (1996).
- Zou, X. D., Sukharev, Y. & Hovmöller, S. *Ultramicroscopy* **52**, 436–444 (1993).

SUPPLEMENTARY INFORMATION. Available on Nature's World-Wide Web Site <http://www.nature.com>. Paper copies are available from Mary Sheehan at the London editorial office of Nature.

ACKNOWLEDGEMENTS. We thank V. Duppel for performing the EDXS analysis. X.D.Z. is supported by an NFR postdoctoral fellowship. S.H. was supported by the Swedish NFR.

CORRESPONDENCE should be addressed to X.D.Z. (e-mail zou@struc.su.se).

Increased activity of northern vegetation inferred from atmospheric CO_2 measurements

C. D. Keeling*, J. F. S. Chin† & T. P. Whorf*

* Scripps Institution of Oceanography, La Jolla, California 92093-0220, USA

† Mauna Loa Observatory, NOAA/CMDL, Hilo, Hawaii 96721, USA

THROUGHOUT the Northern Hemisphere the concentration of atmospheric carbon dioxide rises in winter and declines in summer, mainly in response to the seasonal growth in land vegetation^{1–4}. In the far north the amplitude of the seasonal cycle, peak to trough, is between 15 and 20 parts per million by volume⁵. The annual amplitude diminishes southwards to about 3 p.p.m. near the Equator, owing to the diminishing seasonality of plant activity towards the tropics. In spite of atmospheric mixing processes, enough spatial variability is retained in the seasonal cycle of CO_2 to reveal considerable regional detail in seasonal plant activity⁶. Here we report that the annual amplitude of the seasonal CO_2 cycle has increased by 20%, as measured in Hawaii, and by 40% in the Arctic, since the early 1960s. These increases are accompanied by phase advances of about 7 days during the declining phase of the cycle, suggesting a lengthening of the growing season. In addition, the annual amplitudes show maxima which appear to reflect a sensitivity to global warming episodes that peaked in 1981 and 1990. We propose that the amplitude increases reflect increasing assimilation of CO_2 by land plants in response to climate changes accompanying recent rapid increases in temperature.

Our most detailed record of the seasonal signal in atmospheric CO_2 has been obtained at Mauna Loa Observatory, Hawaii (20° N, 156° W). We determined the annual amplitude from the interannually detrended record by computing a 4-harmonic sea-

sonal function with invariant phasing, but with linearly increasing amplitude with time, as described in ref. 3. We then estimated this amplitude for each calendar year (Fig. 1a, dots in upper plot) by least-square fits of this phase-invariant function to yearly segments of the detrended data. By a second method designed to remove most of the constraint on the computation of amplitude imposed in the first method by a phase-locked signal, we also computed separate 4-harmonic seasonal functions for successive overlapping 3-year segments of the record. The gain of each function was afterwards adjusted to achieve a best fit to the middle year. Annual amplitudes determined by this second method differed only slightly from those determined by the first method. Also, the total gain in amplitude from 1964 to 1994, based on a linear fit to the entire record by the two methods is nearly the same, $20.2 \pm 1.4\%$ and $19.7 \pm 2.0\%$, respectively. Here we show amplitudes determined only by the first method. We have observed a similar, although somewhat larger and more irregular, increase in annual amplitude since 1970 at La Jolla, California (33° N, 117° W)⁷, confirming that the Mauna Loa record is not exceptional for middle latitudes of the Northern Hemisphere.

The annual amplitude observed at Mauna Loa since 1958 shows almost no trend until the mid-1970s¹; then it increased almost steadily until 1984; decreased for several years, and increased again to a record maximum in 1991. By smoothing the amplitudes with a spline fit, a quasi-decadal oscillatory pattern is seen with peaks in 1983 and 1992 (solid curve, Fig. 1a).

Changes in phase of the Mauna Loa seasonal cycle, characterized by the times of downward zero crossing of the successive harmonic functions, determined by the second method described above, show no trend until the mid-1970s (Fig. 1a, lower plot). Thereafter the time of crossing advanced irregularly, becoming about 7 days earlier near the end of the record. Also, the timing of the minimum advanced irregularly after 1980 by about 3 days while the maximum and the upward zero crossing showed no significant long-term changes. These features, taken together, suggest that the growing season of plants has lengthened during approximately the same time period in which the annual amplitude has increased.

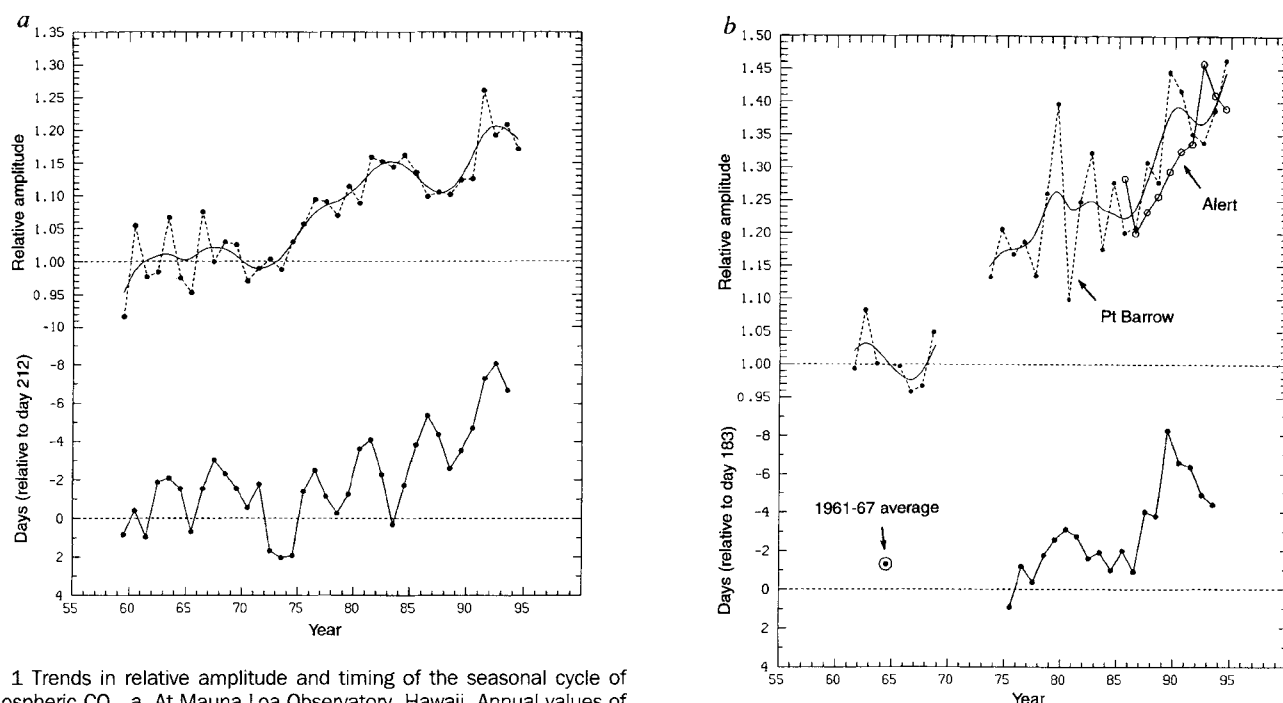


FIG. 1 Trends in relative amplitude and timing of the seasonal cycle of atmospheric CO_2 . *a*, At Mauna Loa Observatory, Hawaii. Annual values of the amplitude (upper plot, dots connected by dashed lines) were determined from weekly averaged continuous concentration data¹⁵ fitted annually to a phase-locked 4-harmonic seasonal function. This function, together with a linearly increasing gain factor, was first established by a fit to the full record³, after which the gain was redetermined separately for each year. The linearly increasing gain factor, referenced to 1964, was computed to be $0.675 \pm 0.045\%$ per year. A smoothing spline¹⁹ (solid curve, with standard error, σ_A , of 0.028%) shows quasi-decadal changes in the relative amplitude. The timing of the downward zero crossing of the seasonal cycle (lower plot, dots connected by solid line segments), was obtained from 3-year fits to the CO_2 record, as described in the text and plotted relative to the 212th day of the year. The zero crossing is with respect to the seasonally adjusted CO_2 concentration determined by the method of ref. 3. *b*, As *a*, but

for Point Barrow, Alaska. The data from 1973 to 1994 are from flask samples collected 2–4 times per month, augmented from 1961 to 1967 by continuous CO_2 measurements⁸, also calibrated by our laboratory. The fitting procedure (which spans a record gap from 1968 to 1973) yields a linear gain factor of $1.318 \pm 0.100\%$ per year, again referenced to 1964. A smoothing spline (σ_A of 0.050%, as in *a*), shows quasi-decadal changes. Also shown is the trend in amplitude at Alert, Canada (connected open circles in upper plot), plotted so that the sum of the squared annual differences from Point Barrow data, is minimized. The zero crossing time at Point Barrow (lower plot) is plotted relative to the 183rd day of the year. The circled dot in 1964 was determined from 3-year fits for 1961–63 and 1965–67, averaged.

In the Arctic an atmospheric CO_2 record nearly as long as that at Mauna Loa exists for Point Barrow, Alaska (71°N , 157°W)^{8,9}. The total gain in annual amplitude from the mean for 1961–67 is 39.5 ± 3.0 , approximately twice that at Mauna Loa. The smoothed amplitudes after 1970 show a quasi-decadal pattern not unlike that for Mauna Loa, but with periods of high amplitude near 1980 and 1990 and a low amplitude near 1985 (Fig. 1*b*, upper plot) thus coming sooner than the corresponding periods for Mauna Loa. In addition, large year-to-year fluctuations in amplitude have occurred in the Point Barrow record.

For another Arctic station, at Alert, Canada (82°N , 62°W), with a record beginning in 1985 (open circles in Fig. 1*b*), the annual amplitude generally shows only small year-to-year fluctuations, but has an upward trend similar to that at Point Barrow. The general occurrence of an upward trend in the Arctic and subarctic since 1985 is further confirmed by our harmonic analysis of observations at Mould Bay, Canada and Cold Bay, Alaska⁷.

At Point Barrow the time of downward zero crossing (Fig. 1*b*, lower plot) has advanced nearly 7 days since 1975, approximately as at Mauna Loa. The pattern of advance, unlike that at Mauna Loa, tends to be quasi-decadal, approximately in phase with the smoothed amplitude. Other features of the seasonal cycle have varied as well, as at Mauna Loa. Mindful of the gap in the record from 1968 to 1973 we have compared the shape of the average cycle for 1961–67 with that for 1988–94. From the first averaging period to the second, the period of low summertime concentrations has broadened, and the succeeding period of rapidly rising CO_2 comes later in the season than before. In early November when formerly the rate of rise slackened off, in recent years it has continued to rise steeply for another month and a half. These

changes, taken together, suggest a progressive lengthening of the growing season, as also suggested by the Mauna Loa record.

We have compared the trend in annual amplitude at Mauna Loa with surface air temperature over land averaged annually between 30°N and 80°N , the zone where most of the seasonal growth of plants takes place in the Northern Hemisphere (see Fig. 3*c* of ref. 6). On the decadal timescale, in which short-term interannual variations are filtered out, the temperature record (upper solid curve of Fig. 2*a*) shows a quasi-decadal pattern with extremes in warmth near 1981 and 1990, about two years before decadal maxima in CO_2 amplitude (lower solid curve). Linear regression of the annual amplitude and temperature data indicate (by lowest χ^2) a lag of between one and two years (see Fig. 2 legend).

We have similarly compared the Point Barrow annual amplitude with land temperatures north of 50°N (Fig. 2*b*), where model simulations⁶ indicate that about two-thirds of the Barrow amplitude signal is derived. On the decadal timescale, the temperature record shows a pattern similar to that north of 30°N , with which the CO_2 amplitude is approximately in phase. However, a linear regression of this data gave the lowest χ^2 for amplitude lagging temperature by one year (see Fig. 2).

The observed correlations of annual CO_2 amplitude with land surface temperature suggest large-scale responses of the carbon cycle to climate change. The nature of these responses, however, is not obvious. Neither the signals nor their correlations and lags with temperature were anticipated from prior knowledge of climate or of the global carbon cycle. Some clues as to the possible causes of these signals will now be discussed. We first address long-term increases and then quasi-decadal fluctuations.

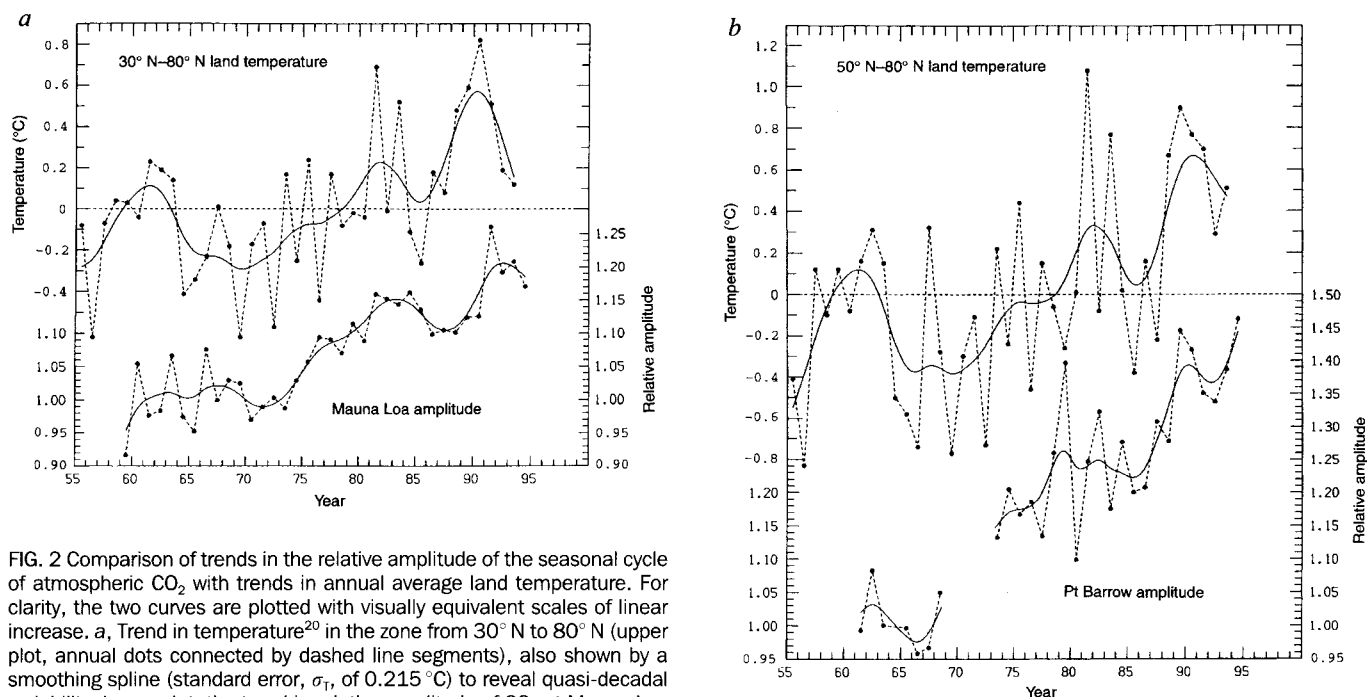


FIG. 2 Comparison of trends in the relative amplitude of the seasonal cycle of atmospheric CO_2 with trends in annual average land temperature. For clarity, the two curves are plotted with visually equivalent scales of linear increase. *a*, Trend in temperature²⁰ in the zone from 30° N to 80° N (upper plot, annual dots connected by dashed line segments), also shown by a smoothing spline (standard error, σ_T , of 0.215 °C) to reveal quasi-decadal variability. Lower plot, the trend in relative amplitude of CO_2 at Mauna Loa Observatory, as in Fig. 1a. The average increase of amplitude with respect to temperature was found by linear regression of annual data, allowing for errors in both variables, represented by σ_A and σ_T , respectively (ref. 21, p. 660). Best-fit estimates based on the lowest χ^2 for the period 1973–94 (for comparison with Point Barrow) showed nearly identical results for lags of one and two years in amplitude relative to temperature, and gave an increase of $19.0 \pm 4.0\%$ per °C (reduced $\chi^2 = 1.2$) for the two-year lag. (If run from 1966 to 1994 for a two-year lag, the result was $22.7 \pm 3.7\%$

per °C.) *b*, Similar comparison of the amplitude at Point Barrow, Alaska, as in *a*, but using the trend in temperature between 50° N and 80° N (σ_T of smoothing spline 0.320 °C). The best-fit regression in this case was for a lag in amplitude of one year and gave an increase of amplitude with temperature for 1973–94 of $25.5 \pm 5.6\%$ per °C (reduced $\chi^2 = 0.9$). Linear correlation coefficients confirmed the lags for both stations.

With respect to long-term changes we focus on the period 1973–94, when we have uninterrupted CO_2 data for both Mauna Loa, Hawaii and Point Barrow, Alaska. Based on linear regressions of annual CO_2 amplitude to temperature over this 21-year period (see Fig. 2 legend), the amplitude at Mauna Loa increased by $19 \pm 4\%$ per °C with respect to land temperatures north of 30° N, and Point Barrow by $26 \pm 6\%$ per °C with respect to temperature north of 50° N. Short-term studies indicate that photosynthesis of land plants typically increases by 40–100% for a 10 °C increase in temperature (see p. 147 of ref. 10), with plant respiration (see p. 149 of ref. 10) and soil respiration¹¹, showing a somewhat larger response. These short-term responses do not approach the high sensitivity to temperature implied by our CO_2 data. Our less-complete data from before 1973 indicate even higher implied sensitivities to temperature.

A few per cent of the overall increase in annual amplitude may be explained by increased seasonality in fossil-fuel combustion and an out-of-phase oceanic cycle, but both effects have been shown by model simulations to be small⁶. Also, the growth of plants is likely to have been stimulated by the 14% increase in ambient CO_2 caused by the rise in global CO_2 concentrations since the beginning of our records, but this effect is predicted^{12,13} to explain no more than about a 6% increase in amplitude in 35 years. Changes in atmospheric transport may have contributed to the observed signals, but the general finding of substantial amplitude increases at stations deliberately placed to minimize local influences, makes it unlikely that our conclusions would have been altered by a more-detailed analysis of atmospheric transport.

Thus, the observed annual amplitude increases, especially the large increase found in the Arctic, does not seem to be adequately explained by either direct, short-term responses of plants to temperature or by other factors unrelated to temperature. We propose that an increase in length of the growing season, suggested by changes in timing of the seasonal cycle of CO_2 , noted above, may be an important cause of amplitude increases. The

finding of Chapman and Walsh¹⁴ that warming at high latitudes has been greatest in winter and spring (up to 4 °C in 30 years) supports this hypothesis.

Any explanation for the recent increases in annual amplitude of CO_2 should also account for its evident correlation with temperature on the decadal timescale. In this connection we note that the abundance of atmospheric CO_2 , also appears to correlate decadal with temperature¹⁵, as though variations in the seasonal cycle of CO_2 have caused imbalances in the annual exchange of plant CO_2 with the atmosphere. We now discuss this possibility.

Interannual variations in atmospheric CO_2 abundance are partly contributed by oceanic processes, but model simulations suggest that the quasi-decadal signal in abundance has been predominantly terrestrial⁹, and therefore should largely reflect imbalances between the uptake of CO_2 by vegetation, expressed as net primary production (NPP), and release of CO_2 from plant litter and soils, caused by heterotrophic respiration (HR). Because the seasonal signal in CO_2 is caused mainly by NPP during summer and by HR during winter, (though always made up of NPP – HR), the phase relationship between the decadal signals in amplitude and in abundance should be related to how interannual variations in NPP and HR have affected both signals.

Over the time period of atmospheric CO_2 measurements, quasi-decadal variations in CO_2 abundance have been observed to be nearly in phase with quasi-decadal variations in global temperature^{9,15} which have nearly the same phasing as the temperatures north of 30° N, shown in Fig. 2. The decadal signal in the annual average net ecosystem flux, NPP – HR, has therefore lagged temperature by a quarter of a cycle, that is, by approximately two years. Because the observed quasi-decadal signal in annual CO_2 amplitude, as shown in Fig. 2a, has also lagged temperature by about two years, decadal maxima in amplitudes have evidently occurred approximately when annual NPP has maximally exceeded annual HR, and *vice versa*.

These phase relations suggest that NPP, on a decadal timescale

has been stimulated when it warms and diminished when it cools, with the response lagging temperature. Heterotrophic respiration, HR, may have also been affected, either directly or in response to the changes in NPP, in a manner to produce a two-year phase lag in the difference, NPP – HR. Variations in mean annual temperature have been too small to be a likely direct cause of such variations, but coherent variations in the length of the growing season, especially at higher latitudes, may well have produced large variations in NPP with consequent variations in HR.

Although the quasi-cyclic variations in the carbon cycle seen in our data would therefore appear to arise from natural variability in temperature, this variability is not obviously the cause of the longer-term increases in annual CO₂ amplitude reported here. These striking increases over 30 years could represent unprecedented changes in the terrestrial biosphere, partially in response to some of the highest global annual average temperatures since the beginning of modern records^{16,17}, and partially in response to plant growth being stimulated by the highest concentrations of atmospheric CO₂ in the past 150,000 years (ref. 18). □

Received 19 April; accepted 20 May 1996.

1. Hall, C. A. S., Ekdahl, C. A. & Wartenberg, D. E. *Nature* **255**, 136–138 (1975).
2. Pearman, G. I. & Hyson, P. J. *geophys. Res.* **85**, 4468–4474 (1980).
3. Bacastow, R. B., Keeling, C. D. & Whorf, T. P. *J. geophys. Res.* **90**, 10529–10540 (1985).

4. D'arrigo, R., Jacoby, G. C. & Fung, I. Y. *Nature* **329**, 321–323 (1987).
5. Conway, T. J. et al. *J. geophys. Res.* **99**, 22831–22855 (1994).
6. Heimann, M., Keeling, C. D. & Tucker, C. J. in *Aspects of Climate Variability in the Pacific and the Western Americas* (ed. Peterson, D. H.) 277–303 (Am. Geophys. Union, Washington DC, 1989).
7. Whorf, T. P. & Keeling, C. D. *Summary Report 1990*, 124–126 (Climate Modeling & Diag. Lab. Vol. 19, NOAA, Boulder, CO, 1991).
8. Kelley, J. J. Jr *Scientific Report*, Contract N00014-67-A-0103-0007 NR 307-252 (US Office of Naval Research, Arctic Program, Washington DC, 1969).
9. Keeling, C. D. et al. in *Aspects of Climate Variability in the Pacific and the Western Americas* (ed. Peterson, D. H.) 165–236 (Am. Geophys. Union, Washington DC, 1989).
10. Larcher, W. *Ökologie der Pflanzen* (Ulmer, Stuttgart, 1984).
11. Raich, J. W. & Schlesinger, W. H. *Tellus* **44B**, 81–99 (1992).
12. Kohlmaier, G. H. et al. *Tellus* **41B**, 487–510 (1989).
13. Luo, Y. & Mooney, H. A. in *Responses of Terrestrial Ecosystems to Elevated CO₂* (eds Koch, G. W. & Mooney, H. A.) 381–397 (Academic, San Diego, in the press).
14. Chapman, W. L. & Walsh, J. E. *Bull. Am. met. Soc.* **74**, 33–47 (1993).
15. Keeling, C. D., Whorf, T. P., Wahlen, M. & van der Plicht, J. *Nature* **375**, 666–670 (1995).
16. Aldous, P. *Nature* **349**, 186 (1991).
17. Jones, P. D. & Briffa, K. R. *Global Surface Air Temperature Variations during the Twentieth Century Part 1. Spatial, Temporal and Seasonal Details, Holocene 2.2*, 165–179 (Climate Res. Unit, Univ. East Anglia, Norwich, 1992).
18. Barnola, J. M., Raynaud, D., Korotkevich, Y. S. & Lorius, C. *Nature* **329**, 408–414 (1987).
19. Reinsch, C. H. *Numerische Mathematik* **10**, 177–183 (1967).
20. Jones, P. D. *J. Clim.* **7**, 1794–1802 (1994).
21. Press, W. H., Teukolsky, S. A., Vetterling, W. T. & Flannery, B. P. *Numerical Recipes in Fortran* 2nd edn (Cambridge Univ. Press, 1992).

ACKNOWLEDGEMENTS. We thank P. Jones for providing monthly gridded surface air temperature data; P. Tans, N. Trivett and their associates for assistance over many years in acquiring atmospheric data at Mauna Loa Observatory, Point Barrow, Alaska, and Alert Station, Canada; and R. Bacastow, R. Hunt, R. Keeling, J. Kelley, S. Piper and J. Walsh for their advice. This work was supported by the US NSF, the US Department of Energy, NASA and NOAA.

CORRESPONDENCE should be addressed to C.D.K. (e-mail cdkeeling@ucsd.edu).

Structural control on sea-floor hydrothermal activity at the TAG active mound

Martin C. Kleinrock* & Susan E. Humphris†

* Department of Geology, Vanderbilt University, Nashville, Tennessee 37235, USA

† Department of Geology and Geophysics, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

THE TAG active mound^{1–13}, on the Mid-Atlantic Ridge near 26° N, is one of the largest known, actively forming volcanogenic massive sulphide deposits. Construction of such a deposit requires that the upflow of hydrothermal fluids be localized at this site over tens of thousands of years, but the cause of this localization has been controversial. Two popular hypotheses propose that it results from young volcanism immediately adjacent to the mound⁶, or from the intersection of a 'transfer fault', orthogonally crossing the rift-valley floor, with ridge-parallel faults⁸. Here we present high-resolution sonar and photographic data which provide no evidence for young eruptions nearby, or for transverse, through-going faults intersecting the TAG mound. Instead, our data suggest that the localization of hydrothermal activity at the TAG mound is strongly controlled by permeable conduits at the intersection of an actively developing ridge-parallel fissure/fault complex with a newly recognized but older, oblique fault system. These observations may aid the interpretation of similar structures in ancient sea-floor sulphide deposits, where the tectonic disruption caused by emplacement on land has obscured the original geometry.

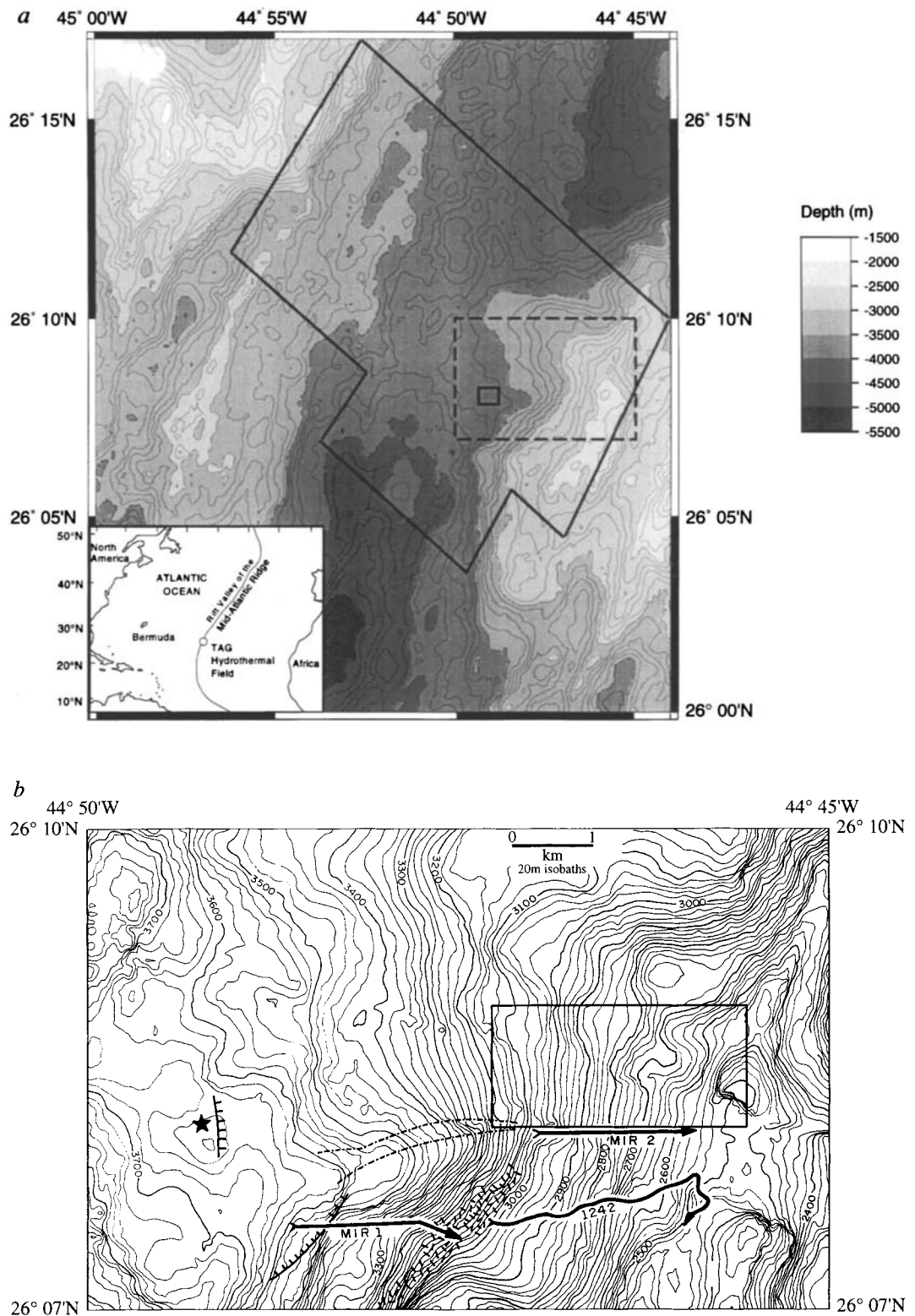
Today, hydrothermal activity at temperatures above 25 °C in the TAG (Trans-Atlantic Geotraverse) area is confined to the ~200-m-diameter active mound, 26° 8.2' N, 44° 49.5' W (Fig. 1). This mound of polymetallic sulphides and massive anhydrite¹¹, located near the Mid-Atlantic Ridge median valley wall ~3 km east of the neovolcanic axis, appears to have been active episodically for at least 20,000 years (refs 7 and 12), and is estimated to contain about 3.9 million tonnes of sulphides¹³. We collected high-

resolution sidescan sonar imagery with co-registered bathymetry of a 10 × 10 km region of the rift valley and detailed photography (including over 30,000 electronic still camera images) from the mound and its environs using the deeply towed ARGO-II and DSL-120 instrument packages^{14–16}.

Based primarily on the 120-kHz sidescan data, we divide the rift valley in this area into four distinct morphological zones (Fig. 2): a neovolcanic zone, an active fissure/fault zone, and eastern and western terraced zones. The TAG active hydrothermal mound (Figs 2 and 3) lies within the active fissure/fault zone at the intersection of rift-parallel N-to-NNE-striking faults and fissures and a set of newly identified, older ENE-striking faults. Within ~50 m to the east of the mound, we observed a 20-m-high westward-dipping fault striking NNE (Fig. 3a). The presence of fresh talus at the base of this fault and other NNE faults within this zone, combined with a progressive eastward increase in fissure/fault length and throw¹⁷, show that this area is actively extending and the rift-parallel NNE structures are still developing.

Bathymetric and sonar images constructed from our new data (Fig. 3) present the first synoptic view of the shape and nature of the TAG active mound. The distinctly circular, 200-m-diameter mound shoals to a depth of 3,630 m, ~50 m above the surrounding sea floor. The mound comprises two well defined platforms, a lower 150-m-diameter platform at 3,650–3,655 m and an upper 90-m-diameter platform at 3,642–3,650 m. The upper platform contains the previously described^{4,10,11} primary black-smoker complex and a NNE-trending depression that was absent during previous studies in 1986–90. The alignment of this depression with NNE-trending faults and fissures just north and south of the mound suggest that the mound is being extended WNW–ESE in concert with its environment. Whereas previous studies found black-smoker venting only at the primary black-smoker complex¹¹, we observe a new and separate suite of black-smoker vents on the upper platform, many along fissures near or within this depression (Fig. 4). These changes suggest recent surficial collapse linked to continuing fissure/fault development, perhaps with associated changes in the subsurface permeability and fluid pathways. The generation of the new depression atop the mound and the coincident development of new black-smoker vents show that contemporaneous tectonic deformation and hydrothermal deposition modified the three-dimensional structure and hydrogeology of the TAG active mound shortly before our survey.

FIG. 1 *a*, Sea Beam bathymetry of the TAG spreading segment²⁴ showing the areas of the DSL-120 sidescan sonar survey (large solid-line polygon) and the ARGO-II survey of the TAG active hydrothermal mound (small solid-line box). The DSL-120 instrument was towed ~100 m above the sea floor and collected 120-kHz sidescan sonar imagery and co-registered bathymetry; it was also equipped with a transmissometer, a magnetometer, and other instruments; the superb resolution and small footprint of the DSL-120 bathymetry allow for the first time the acquisition of highly detailed images such as those of the TAG active mound and environs shown in this Letter. The ARGO-II instrument was towed ~5–15 m off the sea floor and collected electronic still photographs, full-motion video imagery, and 200-kHz sidescan sonar data, and was also equipped with a magnetometer, an Imagenex narrow-beam scanning sonar, and other instruments. The dashed-line box shows the area shown in more detail in *b*. Grey shades change every 500 m; 100-m contours. Inset, location of the TAG segment on the Mid-Atlantic Ridge near 26° N. *b*, Detailed map of a small region of the eastern median valley floor and wall (after ref. 6). The box and the arrows show the regions of the eastern valley wall covered by submersible dives^{25,8,6}, which suggested the presence of an east–west transfer fault⁸. Also shown are two fault line scarps (solid lines, hatches downhill) and an area (dashed) of gabbro outcrops⁶. The star represents the active TAG hydrothermal mound. The east–west feature outlined by the dash-dotted line was interpreted to be an erosional gully⁶.

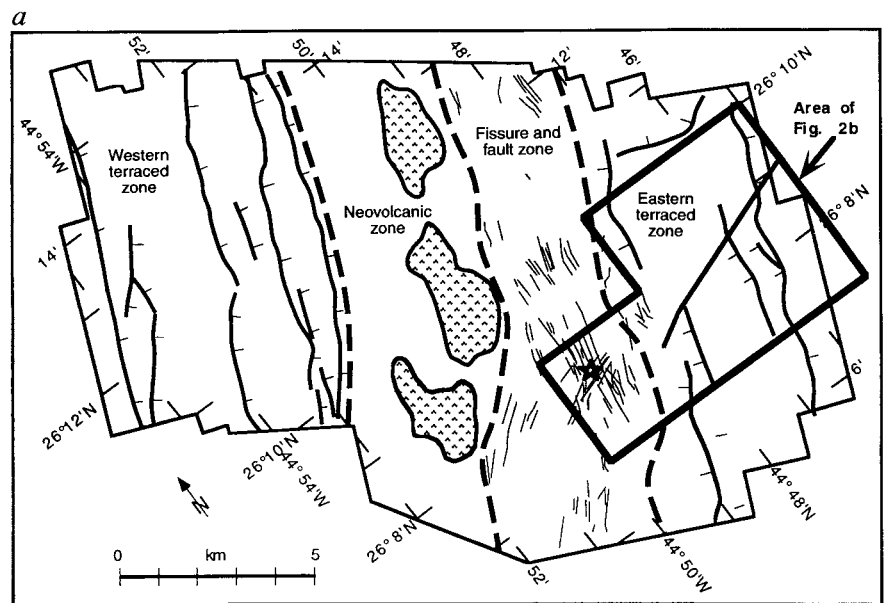


Our high-resolution data set allows us to test the two commonly held hypotheses for the cause of localization of upflow of hydrothermal fluids at the TAG active mound: intersection of a ridge-parallel fault with a cross-axis transfer fault⁸, and a concentrated local heat source⁶.

We detect no evidence for an east–west fault continuing from the rift shoulder and intersecting the rift-valley floor at the TAG active mound, such as that previously hypothesized by Karson and Rona⁸. Based on observations of different attitudes of sedimen-

tary beds high on the rift-valley wall, Karson and Rona⁸ suggested that an east–west fault runs along an erosional gully (Fig. 1*b*) and represents an accommodation zone separating two independent fault blocks. Because the TAG active mound lies along-strike, they suggested that the intersection of this lineament with ridge-parallel faults may localize hydrothermal upflow at the mound. Our detailed mapping now provides complete coverage of the region to the east of the mound, including the valley wall. Although there is evidence for an east–west fault on the rift-valley

FIG. 2 a, Simplified geological interpretation of the $\sim 10 \text{ km} \times 20 \text{ km}$ region of the TAG rift valley based on our new data. Four distinct morphological zones separated by heavy dashed lines have been identified and discussed by Kleinrock and Humphris²⁶. From west to east, they are a zone of heavily sedimented uplifted terraces bounded by distinct rift-parallel faults (heavy lines, hatches downhill) west of the inner valley floor, a 4-km-wide neovolcanic zone comprising a discontinuous series of young unfissured volcanic regions (stippled) in the middle of the rift-valley floor, a zone of closely spaced and actively developing fissures and small faults (thin lines) just east of the neovolcanic zone, and another zone of fault-bounded sedimented terraces further to the east. East of the neovolcanic zone, total extension increases to the east and becomes progressively concentrated on selected fissures and faults. The southern portion of the actively fissuring zone contains an additional set of faults, with an ENE trend, that are cut by the NNE-trending faults and fissures. The TAG active hydrothermal mound (starred) lies within the active fissure/fault zone where the ENE trend intersects the NNE trend and is at least 3 km from any very recent volcanism. b, DSL-120 sidescan sonar imagery of the area of Fig. 1b, with inset showing simplified interpretation of selected features. These data show the pattern of ENE and NNE intersecting faults in the area of the TAG active mound, as well as the larger-scale structures in the vicinity. Note that the east-west transfer fault



inferred by Karson and Rona⁸ does not extend from the area of their *Alvin* dives to the TAG active mound; the scarp associated with the tilt changes they observed is clear on the image, but curves to the south and projects south of the TAG active mound. Also note the presence of throughgoing NNE faults east of the mound showing that there is no major east-west structure extending from rift-valley wall to the mound.



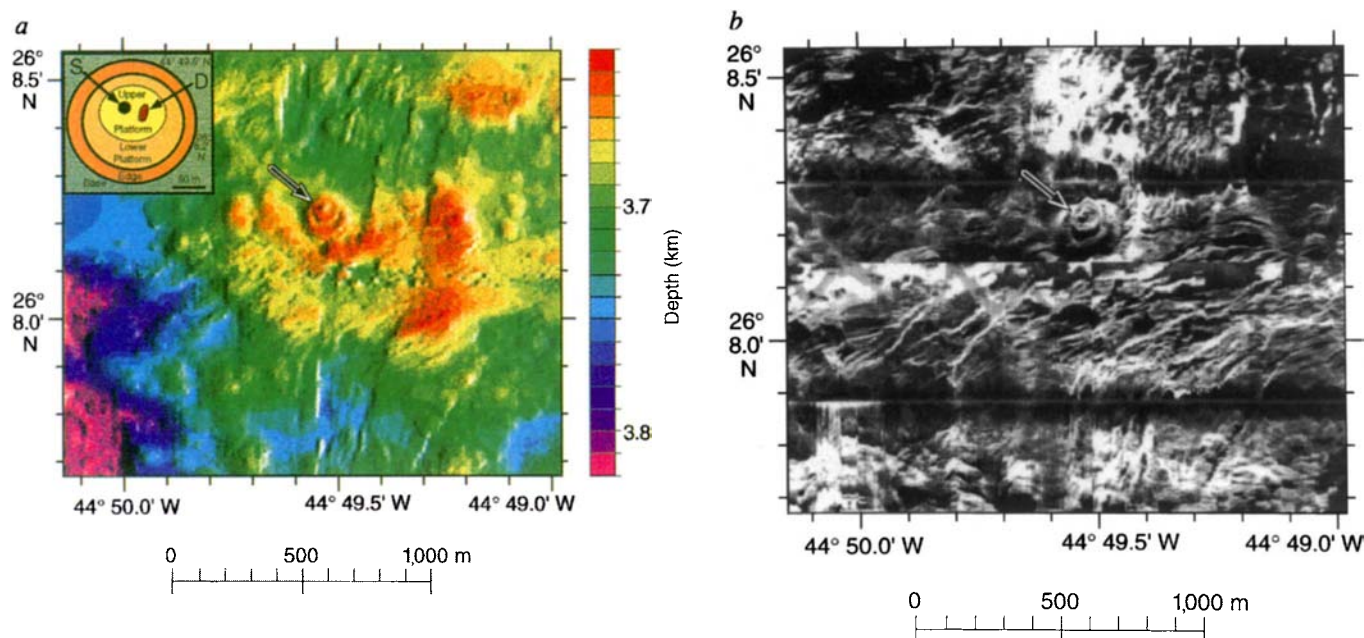


FIG. 3 Colour-coded shaded relief bathymetry (a) and sidescan sonar imagery (b) of $\sim 4 \text{ km}^2$ area including the TAG active mound based on data collected with the DSL-120 near-bottom towed instrument package. In a, reds are shallowest, purples are deepest, and illumination is from the east. The sidescan image is lighter where returns are strongest (rougher sea floor or scarps facing towards the sound source/receiver pair) and black in areas shadowed from the acoustic energy; navigation, slant-range, speed and noise corrections have been applied. The size of the sea floor represented by individual data samples (footprint) measured by the DSL-120 is $\sim 1 \text{ m}^2$, and the vertical resolution is better than 1 m, making these data several orders of magnitude more precise than data collected with traditional multibeam systems¹⁵ (for example, Sea Beam has a $\sim 25,000 \text{ m}^2$ footprint). The 200-

m-diameter circular active mound (arrowed) includes two platforms. In a, inset shows sketch of active mound. The lower platform contains in its southeastern portion the previously identified white smokers of the 'Kremlin' area, so-called for its chimneys with minaret-like tops^{4,10,11}. The upper platform contains the primary black smoker complex ('S'), a 5–20-m-diameter, 10-m-high cone topped by many chimneys vigorously discharging hot (up to 363°C) fluids^{4,10,11}. Also on the upper platform, east of this black smoker complex, lies a newly formed, 10 m deep, NNE-trending, $25 \text{ m} \times 20 \text{ m}$ depression ('D') that has associated with it a new set of fissures and high-temperature vents (see Fig. 4). These images also show faults near the mound.

wall, confirming Karson and Rona's interpretation of a transfer fault separating fault blocks, this fault curves to the southwest and does not continue to the active mound (Fig. 2). Therefore, although we cannot rule out the possibility of effects from broad and complex accommodation zone, it appears that a transfer fault is not responsible for focusing fluid flow at the active mound.

Our data do, however, lead us to the same basic conclusion that the activity at the TAG mound is localized by the intersection of tectonic structures. Figures 2 and 3 show that the mound lies at the intersection of a zone of ridge-parallel fissures and with a series of obliquely oriented (ENE) faults. We infer that the ridge-parallel structural system is actively extending, based on analysis of the ARGO-II photography that suggests a monotonic increase in strain across the set of ridge-parallel structures to the east and shows fresh fissures cutting the sedimentary cover in the area. Cross-cutting relationships indicate that the oblique ENE faults predate the ridge-parallel structures, suggesting that they were created in an earlier stress regime in which oblique stresses were present, perhaps associated with the terminus of an old fourth-order segment or even as part of an accommodation zone, though further study is required to determine the origin of the early ENE faults.

The actively extending ridge-parallel structures may provide pathways for the upflow of hydrothermal fluids from a heat source at depth and/or to the west under the neovolcanic zone. Brecciation at the points of intersection between these NNE and ENE faults can provide an exceptionally permeable conduit to focus fluid flow. Outflow from a conduit with such a narrow linear geometry would probably generate an axisymmetric sulphide deposit, hence explaining the striking circularity of the mound.

It has also been hypothesized that the hydrothermal activity may be localized to an area of very recent volcanic activity⁶.

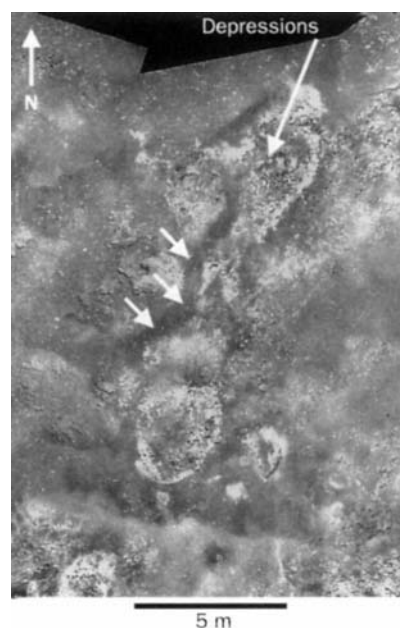


FIG. 4 Electronic still camera digital photomosaic showing black-smoker fluid emanating from a newly developed fissure along the upper platform near the recently formed depression. Previous studies had observed such high-temperature fluids venting only from the primary black-smoker complex, suggesting a recent change in the hydrothermal system occurred before the ODP drilling that postdated our survey. Image is $13 \text{ m} \times 24 \text{ m}$. Locations of venting black-smoker fluids are identified with arrows.

Evaluation of volcanic construct ages, based primarily on sediment distribution inferred from near-bottom photography, indicates that all volcanics within 3 km of the mound have at least a thin covering of pelagic sediments and are hence probably more than a few hundred years old (using a sedimentation rate⁶ of 2 cm kyr^{-1}); the vast majority are heavily sedimented and must have erupted thousands of years ago. Volcanic domes described previously in this part of the rift valley¹⁰ have experienced no active extrusion in the recent past. Younger volcanics are only found 3 km to the west in the neovolcanic zone (Fig. 2). Using reasonable estimates (specific heat of $1 \text{ kJ}^{-1} \text{ K}^{-1}$, latent heat of fusion of 400 kJ kg^{-1} , density of $3,000 \text{ kg m}^{-3}$), the heat produced in solidifying and cooling from $1,200$ to 200°C a 1-km^3 magma body (reasonable for a body feeding a volcano on the Mid-Atlantic Ridge¹⁸) is $\sim 4 \times 10^{18} \text{ J}$; an energy balance with the estimated $500\text{--}940 \text{ MW}$ of total heat flux from the TAG active mound¹⁹ suggests that a single such magma body can maintain fluid circulation for ~ 200 years. This estimation and the 50-year age for the present venting episode^{7,12} require either a new, or very recently replenished, heat source. Lacking extrusive evidence of any nearby heat source, we infer that either the heat presently driving the circulation comes from intrusions lacking associated extrusion^{6,10}, or circulation along deep fracture systems allows water to carry heat several kilometres from its source near the neovolcanic zone to the present venting site^{20–22}. If the latter is true, this strongly suggests concentration of hydrothermal fluid flow along faults (for example, the aforementioned fault immediately east of the mound) dipping from the edge of the rift-valley floor into the neovolcanic zone²². Regardless of the nature of the heat source, we have documented here for the first time the presence and true nature of faults controlling the localization of hydrothermal activity at the TAG active mound. □

Received 26 April; accepted 29 May 1996.

- Rona, P. A., Klunkhammer, G., Nelsen, T. A., Trefry, J. H. & Elderfield, H. *Nature* **321**, 33–37 (1986).
- Eberhart, G. L., Rona, P. A. & Honnorez, J. *Mar. geophys. Res.* **10**, 233–259 (1988).
- Campbell, A. C. et al. *Nature* **335**, 514–519 (1988).
- Thompson, G., Humphris, S. E., Schroeder, B., Sulanowska, M. & Rona, P. A. *Can. Mineralogist* **26**, 697–711 (1988).
- Lisitsyn, A. P., Bogdanov, Y. A., Zonenshain, L. P., Kuz'min, M. I. & Sagalevich, A. M. *Int. Geol. Rev.* **31**, 1183–1198 (1989).
- Zonenshain, L. P., Kuzmin, M. I., Lisitsyn, A. P., Bogdanov, Y. A. & Baranov, B. V. *Tectonophysics* **159**, 1–23 (1989).
- Lalou, C. et al. *Earth planet. Sci. Lett.* **97**, 113–128 (1990).
- Karson, J. A. & Rona, P. A. *Geol. Soc. Am. Bull.* **102**, 1635–1645 (1990).
- Tivey, M. A., Rona, P. A. & Schouten, H. *Earth planet. Sci. Lett.* **115**, 101–115 (1993).
- Rona, P. A. et al. *Econ. Geol.* **18**, 1989–2017 (1993).
- Tivey, M. K., Humphris, S. E., Thompson, G., Hannington, M. D. & Rona, P. A. *J. geophys. Res.* **100**, 12527–12555 (1995).
- Lalou, C. et al. *J. geophys. Res.* **98**, 9705–9713 (1993).
- Humphris, S. E. et al. *Nature* **377**, 713–716 (1995).
- Ballard, R. D., Yoerger, D. R., Stewart, W. K. & Bowen, A. in *Proc. IEEE Oceans Conf.* 71–75 (IEEE, 1991).
- Kleinrock, M. C. in *CRC Handbook of Geophysical Exploration at Sea* (ed. Geyer, R.) 35–86 (CRC, Boca Raton, 1992).
- Stewart, W. K., Chu, D., Malik, S., Lerner, S. & Singh, H. *IEEE J. Oceanic Engng.* **19**, 599–610 (1994).
- Shaw, P. A., Kleinrock, M. C. & Humphris, S. E. (abstr.) *Eos* **75**, 649 (1994).
- Smith, D. K. & Cann, J. R. *J. geophys. Res.* **365**, 707–715 (1993).
- Rudnicki, M. D. & Elderfield, H. *J. Volcan. geotherm. Res.* **50**, 161–172 (1992).
- Scott, M. R., Scott, R. B., Rona, P. A., Butler, L. W. & Nalwalk, A. *J. Geophys. Res. Lett.* **1**, 355–358 (1974).
- Temple, D. G., Scott, R. B. & Rona, P. A. *J. geophys. Res.* **84**, 7453–7466 (1979).
- Thompson, G., Mottl, M. J. & Rona, P. A. *Chem. Geol.* **49**, 243–257 (1985).
- Macdonald, K. C. et al. *Nature* **335**, 217–225 (1988).
- Purdy, G. M., Sempere, J.-C., Schouten, H., Dubois, D. L. & Goldsmith, R. *Mar. geophys. Res.* **12**, 247–252 (1990).
- Rona, P. A. et al. *J. geophys. Res.* **89**, 11365–1377 (1984).
- Kleinrock, M. C. & Humphris, S. E. *Geophys. Res. Lett.* (in the press).

ACKNOWLEDGEMENTS. We thank the officers and crew of the RV *Knorr* for their fine work during a difficult survey, the Woods Hole Oceanographic Institution's Deep Submergence Operations Group for developing and operating the technologically advanced instruments making this type of study possible, the rest of the science party (C. Davis, J. Goff, L. Goldstein, S. Hussensoeder, K. Koga, K. Nakamura, M. Rashid, L. Reiser Wetzel, P. Shaw, M. Sulanowska and C. Van Dover) for their hard work at sea and valuable scientific discussions, and G. Thompson for his guidance from ashore. The help of J. Howland, T. Larrieu, S. Lerner, M. Marra, D. Pardee, M. Rashid, L. Reiser Wetzel, P. Shaw and M. Sulanowska in post-cruise data processing and analysis is gratefully acknowledged. This manuscript was improved by the reviews of R. Haymon and R. Zierenberg. This work was supported by the US NSF.

CORRESPONDENCE should be addressed to M.C.K. (e-mail: kleinrock@vanderbilt.edu).

Effects of gamete traits on fertilization in the sea and the evolution of sexual dimorphism

Don R. Levitan

Department of Biological Science, Florida State University, Tallahassee, Florida 32306-2043, USA

THE evolution of egg and sperm^{1,2}, and more derived forms of sexual dimorphism, is thought to be driven by sperm competition and postzygotic survival; males are limited by fertilizations, females by resources³. Evidence of sperm competition comes from internal fertilizers, or cases where sperm are deposited on eggs⁴, but in free-spawners, the ancestral mating strategy (refs 1,5 but see ref. 6), females are often sperm limited^{7,8}. Laboratory experiments on sea urchins demonstrate that intraspecific differences in gamete attributes, such as egg size, can influence rates of fertilization. Field experiments in which gametes are released and recaptured demonstrate that the influence of gamete traits on fertilization is not overwhelmed by sea conditions, and that variation in gamete traits can have important fitness consequences. These results suggest a new mechanism for the evolution of anisogamy and sexual dimorphism, in which sperm limitation is important, and natural selection for enhanced fertilization acts on females as well as males.

Given that sperm limitation is ubiquitous among free-spawners⁷, it is possible that variation in egg traits influences female fertilization success⁹. Laboratory experiments on the sea urchin *Strongylocentrotus franciscanus* indicate that male–female pairs differ 25-fold in the amount of sperm needed to fertilize 50% of eggs (the f_{50} value). Although there is seasonal variation in gamete quality in echinoids¹⁰, most of the variation in f_{50} can be explained by the combined effects of mean egg size and the quality of sperm (Table 1).

In a second experiment, the size of an egg influenced its probability of fertilization, even within a single clutch. When sperm were limiting, larger eggs were fertilized preferentially (Fig. 1).

Enhanced fertilization may be a result of more frequent sperm collisions with larger eggs⁹, or arise from other aspects of egg quality, such as age, number, or distribution of sperm-receptor sites, that may covary with egg size. The correlation of larger egg sizes with higher fertilization rates has now been documented among *Strongylocentrotus* species⁹, within species, and within individual females.

Because turbulent water movement can reduce the probability of fertilization to almost zero, seconds after gamete release¹¹, the influence of gamete traits on fertilization may be swamped by environmental conditions, and so the fitness consequences of variation in gamete traits might be minimal. Paired laboratory and field experiments performed in Barkely Sound, British Columbia, Canada, tested whether individual differences in fertilization could be detected under field conditions. Laboratory tests determined the f_{50} of gametes from a single male and female. Additional samples from the same male and female spawn were released into the ocean (3–8 m depth) under natural conditions. Sperm were released first, with eggs being released into the sperm plume after a predetermined sperm dispersal time. After two minutes, the free-drifting eggs were recaptured and inspected for evidence of fertilization (Fig. 2a). Daily estimates of turbulent mixing were calculated from video images of the change in area (height and width) of the sperm cloud. Current velocity of the water at the height of gamete release ranged from 0 to 85 cm s^{-1} during the experiments (sampled at 0.5-s intervals). As the sperm dispersal time increased, average efficiency of fertilization decreased (Fig. 2a). However, dispersal time accounted for only

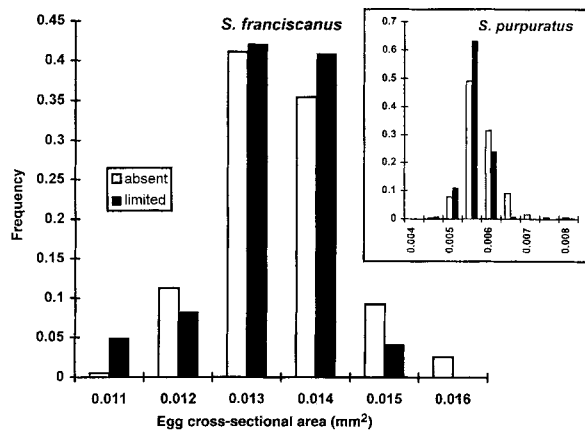


FIG. 1 Size distribution of unfertilized eggs within a single female, when sperm were absent or limiting. In a subset of the laboratory experiments detailed in Table 1, unfertilized eggs were measured that were randomly selected from experimental vials that had approximately 50% of eggs fertilized, and from vials, from the same female, that had no eggs fertilized ($N = 3$ females). When approximately 50% of eggs were fertilized (limiting), the distribution shifted towards smaller eggs than when none were fertilized (absent) ($G = 18.9$, $P < 0.001$). This result indicates that, when sperm are present but not concentrated enough to fertilize all eggs (limiting), larger eggs are preferentially fertilized. At higher sperm concentrations, all eggs, regardless of size, were fertilized (not shown), indicating that smaller eggs can be fertilized if sperm are abundant. Inset, a second species, *S. purpuratus*, demonstrated a similar relation of larger eggs being preferentially fertilized under conditions of sperm limitation ($G = 65.6$, $P < 0.001$, $N = 3$).

21% of the variation in fertilization efficiency, so the remainder must be attributable to fluctuating environmental conditions or gamete quality.

An overall field measure of gamete performance for a particular male–female pair on a particular day is the sperm dispersal time at which 50% of the eggs are fertilized (the t_{50} value). Laboratory gamete performance (f_{50}) explained 55% of the variation in field fertilization rates (t_{50} ; Fig. 2b); gametes that performed well in the laboratory also performed well in the field. A multiple regression incorporating both the sperm diffusion data and the f_{50} data explained 61% of the daily variance in t_{50} .

These results provide at least two insights into the evolution of gamete traits. First, selection on gametes for enhanced fertilization success will be intense not only for males, but also for females. Second, the evolution of anisogamy and gender have been influenced by sperm limitation and selection for enhanced fertilization, rather than simply by sperm competition and postzygotic survivorship.

Theories of the evolution of anisogamy and optimal egg size invoke postzygotic factors as the selective agent on egg size^{1,2,12,13}. This conclusion is reasonable if there is little variance in female fertilization success. Increasing evidence indicates that female fertilization success varies greatly⁷, and my results indicate that, in free-spawning animals, egg size may be under intense selection for higher rates of fertilization.

Because egg size is an indication of maternal investment, selection on egg size is a function of maternal fitness. Assuming a constant total investment in reproduction, selection for producing a maximum number of eggs will result in vanishingly small eggs, but because lower egg size can decrease both fertilization and postzygotic survivorship^{12,13}, optimizing selection acts to produce an egg of intermediate size. My data support the hypothesis that sperm limitation can shift this optimal egg size dramatically^{9,11} (Fig. 3).

All gender differences in morphology, physiology and behaviour stem from the evolutionary transition from isogamy to

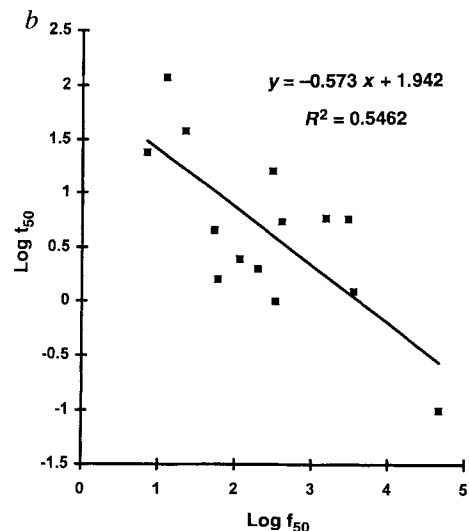
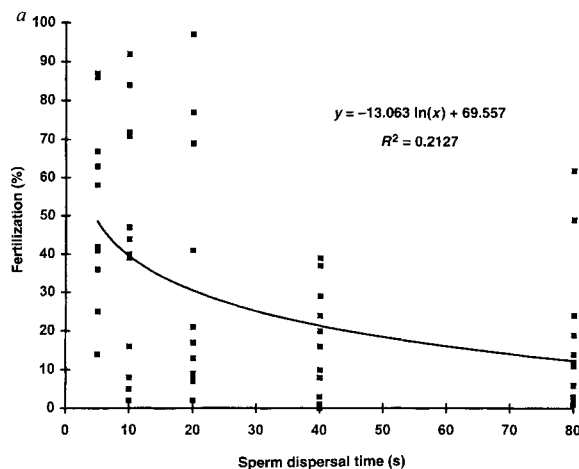


FIG. 2 a, Percentage of eggs fertilized as a function of sperm dispersal time. Dry sperm were diluted with filtered sea water mixed with fluorescein dye. Within 1 min of dilution, 5 ml of the sperm–dye solution was released upwards out of a syringe from a height of 0.5 m at a rate of 1 ml s^{-1} (1×10^7 sperm s^{-1}). The mean spawning rate of *S. franciscanus* injected with KCl was 1.12×10^7 sperm s^{-1} . After a predetermined sperm-dispersal time, $\sim 3 \times 10^7$ eggs mixed with dye were released into the centre of the sperm cloud. At 2 min after release, the free-drifting eggs were collected using a submersible pump, and water was sucked through a filter of mesh size $30 \mu\text{m}$ (ref. 17). The pump was run for 30 s within the gamete cloud and then rinsed out of the cloud for 60 s. Controls indicated that, after the 2-min

sperm–egg interaction time, sperm were too diffuse to result in further fertilization, and all fertilization occurred *in situ* rather than in the filter chamber. Laboratory controls also indicated that the concentration of fluorescein used did not influence rates of fertilization. b, Time of sperm dispersal (t_{50}) in which 50% of eggs are fertilized in the field as a function of fertilizations in the laboratory f_{50} ($P = 0.00025$; regression remains significant at $P = 0.039$, with the bottom-right data point removed). A high t_{50} or a low f_{50} indicates high-performance gametes. Each data point represents a different day and environmental condition ($N = 14$). On a single day the spawn of one male and female were used for both laboratory and field experiments. Individual urchins were used only once.

TABLE 1 Effect of mean egg size and sperm quality on fertilization

Source of variation	d.f.	s.s.	m.s.	F	P
Egg size	1	5.215	5.215	16.17	0.0011
Male	7	5.255	0.751	2.33	0.0803
Residual	15	4.837	0.322		
Corrected total	24	15.307			

Parameter	Estimate	s.e.	P (parameter = 0)
Intercept	-27.216	7.262	0.0019
Egg size	-11.065	2.546	0.0006

$R^2 = 0.684$

Analysis of covariance testing the effect of male quality (main effect) and log egg cross-sectional area (covariate) on the log amount of sperm needed to achieve 50% fertilization (f_{50}). Females (24) were paired with one of eight males (2–4 females per male). Each female was only used once. Dry concentrated sperm from a single male was run through eight 10-fold dilutions, and sperm solution (1 ml) was added to a vial with 9 ml filtered sea water containing ~5,000 eggs from a particular female. Percentage fertilization was assayed after 3 h. The f_{50} was calculated for each male–female pair by fitting the fertilization data from the eight serial dilutions to a fertilization-kinetics model¹⁶ (Fig. 3). Abbreviations: d.f., degrees of freedom; s.s., sum of squares; m.s., mean-square; s.e., standard error.

anisogamy. Previous theories have invoked sperm competition and postzygotic fitness as the selective mechanisms driving the evolution of anisogamy^{1,2}. Divergent selection on isogamous gametes favoured individuals that produced many small gametes (proto-sperm) able to outcompete other individuals for fertilizations when groups of individuals spawn simultaneously. At the other extreme, selection favoured individuals that produced large proto-ova that would optimize postzygotic success when fertilized by a proto-sperm.

Under the sperm-limitation hypothesis, divergent selection on isogamous gametes favoured individuals that produced many small gametes (proto-sperm), increasing the chance of finding

another gamete in a diffuse medium, rather than the chance of outcompeting other sperm (see ref. 15 for a similar, but group-selectionist, argument). Individuals producing large gametes (proto-ova) were selected to increase the probability of fertilization and also postzygotic survivorship. Although both theories predict divergent gamete sizes, they differ in the mechanism driving anisogamy. In the extreme, sperm-limitation theory suggests that sperm competition is a derived condition, resulting from selection for internal fertilization as an escape from sperm limitation. Alternately, sperm limitation and sperm competition represent two ends of a continuum that may vary among taxa, habitats and spawning conditions⁸. Evidence supporting sperm limitation rather than sperm competition is likely to be found in female life-history characters, because both theories predict selection for many small sperm. Female characteristics that may have evolved to increase the probability of fertilization include adaptations in behaviour (aggregation and synchrony), morphology (egg size and structures on eggs, such as jelly coats or accessory cells, that may increase sperm collisions), and physiology (sperm chemoattractants)⁷. With the exception of egg size, these adaptations are independent of postzygotic survivorship, and would not have evolved under sperm-abundant (competitive) conditions. The sperm-limitation hypothesis does not obviate the importance of postzygotic factors, but allows for a more complete understanding of gamete evolution and the evolutionary trajectories of mating systems that later evolved from a free-spawning strategy. □

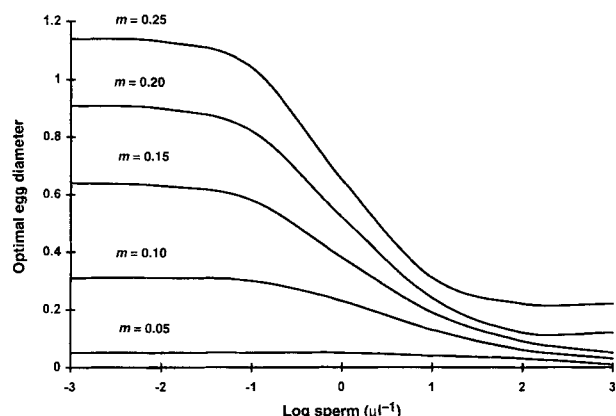


FIG. 3 The optimal egg size, which maximizes the number of successfully settling offspring in a clutch of eggs, varies as a function of ambient sperm concentration and the daily postzygotic larval mortality (m) (ref. 14). Under sperm limitation, optimal egg size increases and becomes more sensitive to changes in postzygotic mortality. The number of settling offspring is calculated as follows. Egg number is determined by $1,000/\text{egg volume}$ (mm^3), which is the number of eggs produced by 1 ml of egg material. Zygote number is the product of egg number, and the proportion of eggs fertilized (ϕ_∞) calculated from a fertilization-kinetics model¹⁶ based on sperm (S_0 ; sperm per microlitre) and egg (E_0 ; $0.01 \mu\text{l}^{-1}$) concentration, sperm–egg interaction time (τ ; 10 min), the sperm–egg collision rate (β_0 ($\text{mm}^3 \text{s}^{-1}$); the product of egg cross-sectional area (mm^2) and sperm velocity (0.130 mm s^{-1} ; ref. 9)), and the fertilization rate constant (β ($\text{mm}^3 \text{s}^{-1}$); 0.0000952 ; ref. 9): $\phi_\infty = 1 - \exp(-\beta S_0 / \beta_0 E_0 (1 - e^{-\beta_0 E_0 \tau}))$. The number of settling zygotes (N_s) is calculated from the number of zygotes produced (N_z) and larval mortality: $N_s = N_z e^{(-tm)}$ (modified from ref. 12), where the larval mortality is calculated by the length of the larval period (t (days) = $18.987 \text{ egg volume} (\text{mm}^3)^{-0.1156}$; ref. 14) and the daily mortality (m ; range based on empirical range noted for the genus *Strongylocentrotus*; ref. 18). The plot represents the egg size with the peak (optimal) number of settling offspring for each combination of larval mortality and sperm concentration.

Received 3 October 1995; accepted 23 May 1996.

- Parker, G. A. in *Sperm Competition and the Evolution of Animal Mating Systems* (ed. Smith, R. L.) 1–59 (Academic, New York, 1984).
- Maynard Smith, J. & Szathmari, E. *The Major Transitions in Evolution* (W. H. Freeman, Oxford, 1995).
- Bateman, A. J. *Heredity* **2**, 349–368 (1948).
- Arnold, S. J. *Am. Nat.* **144**, S126–S149 (1994).
- Wray, G. A. in *Ecology of Marine Invertebrate Larvae* (ed. McEdward, L.) 412–448 (CRC, Boca Raton, FL, 1995).
- Rouse, R. & Fitzhugh, K. *Zool. Scripta* **23**, 271–312 (1994).
- Levitán, D. R. in *Ecology of Marine Invertebrate Larvae* (ed. McEdward, L.) 123–156 (CRC, Boca Raton, FL, 1995).
- Levitán, D. R. & Petersen, C. *Trends Ecol. Evol.* **10**, 228–231 (1995).
- Levitán, D. R. *Am. Nat.* **141**, 517–536 (1993).
- Benzie, J. A. H. & Dixie, P. *Biol. Bull.* **186**, 153–167 (1994).
- Denny, M. W. & Shibata, M. F. *Am. Nat.* **134**, 859–889 (1989).
- Vance, R. R. *Am. Nat.* **107**, 339–352 (1973).
- Smith, C. C. & Fretwell, S. D. *Am. Nat.* **108**, 499–506 (1974).
- Levitán, D. R. *Am. Nat.* **148**, 174–188 (1996).
- Scudo, F. M. *Evolution* **21**, 285–291 (1967).
- Vogel, H., Czhak, G., Chang, P. & Wolf, W. *Math. Biosci.* **58**, 189–216 (1982).
- Mundy, C. N. et al. *Biol. Bull.* **186**, 168–171 (1994).
- Rumrill, S. S. *Ophelia* **32**, 163–198 (1990).

ACKNOWLEDGEMENTS. I thank H. Brook, J. Dalby, D. Levitan and A. Netting for laboratory and field assistance, and T. McGovern, D. Simberloff and J. Travis for comments. This work was supported by Bamfield Marine Station and the NSF.

CORRESPONDENCE and requests for materials should be addressed to the author (e-mail: levitan@bio.fsu.edu).

Trichromatic colour vision in New World monkeys

Gerald H. Jacobs*, Maureen Neitz†, Jess F. Deegan‡ & Jay Neitz‡

* Neuroscience Research Institute and Department of Psychology, University of California, Santa Barbara, California 93106, USA

† Departments of Ophthalmology and Cellular Biology, Medical College of Wisconsin, Milwaukee, Wisconsin 53226, USA

‡ Department of Psychology, California State University, Bakersfield, California 93311, USA

TRICHROMATIC colour vision depends on the presence of three types of cone photopigment. Trichromacy is the norm for all Old World monkeys, apes and humans, but in several genera of New World monkeys, colour vision is strikingly polymorphic¹. The difference in colour vision between these New and Old World primates results from differing arrangements of the pigment genes on the X chromosome. In Old World primates the three photopigments required for routine trichromatic colour vision are encoded by two or more X-chromosome pigment genes and an autosomal pigment gene. New World monkeys typically have only one X-chromosome pigment gene; multiple alleles allow different types of dichromatic colour vision and, in females heterozygous at this locus, variant forms of trichromatic colour vision. Here we report that multiple X-chromosome pigment genes and trichromatic colour vision are the norm for one genus of platyrrhine monkey, the howler monkey, *Alouatta*.

As a part of a survey of cone photopigments in New World monkeys we used a non-invasive electrophysiological index, the electroretinogram (ERG), to determine the number of types of cone photopigment and their spectral properties in howler monkeys (*Alouatta*). We first measured spectral sensitivity functions under test conditions designed to isolate activity from middle- and long-wavelength cones. The results for two females, one black (*A. caraya*) and one red (*A. seniculus*) howler monkey, and for one black male howler monkey, are shown at the top of Fig. 1. We found that the spectral sensitivity functions are too broad to reflect the sensitivity of a single cone, and that there are no obvious differences between the three monkeys. Spectral sensitivity measured in this way reflects the combined signals from middle- and long-wavelength cone pigments², and was essentially identical for these female and male howler monkeys.

An explicit test to determine whether these monkeys had more than one type of X-chromosome-encoded photopigment involved a search for differential effects of chromatic adaptation. Green and red monochromatic lights (540 and 630 nm) flickering in alternation were adjusted in relative intensity until they produced ERGs that were equal in amplitude (a flicker-photometric equation.) These equations were made in the presence of two separate adaptation lights, at 540 and 630 nm. If the retina contains more than a single type of photopigment sensitive to lights from this portion of the spectrum, the equations determined under the two different adaptations should differ, but if there is only a single photopigment the equation value will not change for the two adaptation conditions. Our data (Fig. 1 inset) demonstrates the differences in the equations determined under the two adaptation conditions for samples of dichromatic and trichromatic human subjects, all tested in the same way as the howler monkeys. Note that dichromatic subjects have only a single photopigment sensitive to lights drawn from this part of the spectrum, and show no significant change in the equation values obtained for the two adaptation conditions. Conversely, the human trichromats have two X-chromosome-encoded photopigments, and have significant differences in their equations. The results for the three howler monkeys were very similar to one another, and to the results of the human trichromats.

These measurements of cone sensitivity in howler monkeys contrast sharply with previous findings on New World monkeys. Among animals drawn from species representing nine genera, no single male monkey (from a sample >100 individuals) has previously been found to have more than one X-chromosome encoded photopigment²⁻⁶. This group includes at least one species from each of the six subfamilies of *Platyrrhini*⁷. However, our results strongly imply that howler monkeys have routine trichromatic colour vision (Fig. 1), probably much like that of the catarrhine primates. To investigate the genetic basis of their trichromacy, we examined the X-chromosome photopigment genes of both red and black howler monkeys.

The spectral tuning of primate photopigments is determined by a small number of amino-acid substitutions in the transmembrane portion of the pigment molecule⁸⁻¹². Changes in the nucleotide sequence in exon 5 of the X-chromosome pigment genes are the main reason for the substantial spectral separation of the middle- and long-wavelength pigments underlying trichromacy in many

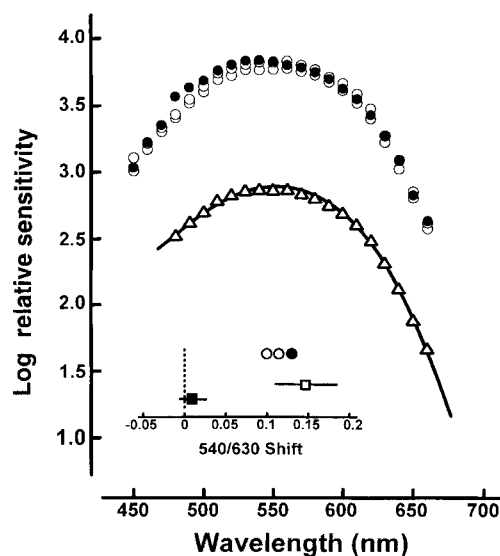


FIG. 1 ERG evaluations of the X-chromosome-linked photopigments in howler monkeys. Top, spectral sensitivity values obtained for three howler monkeys (females, open circles; male, filled circles). Middle, triangles represent average spectral sensitivity data for the three monkeys. The continuous line is the best-fitting summative combination of absorption curves for the two putative cone photopigments of the howler monkey; these two pigments have peak values of 562 and 530 nm. The relative proportions of the two curves were varied in steps of 1% to determine the best fit. The best fit was obtained for a 562/530 weighting of 2.22. To minimize the effect of absorption by potential pre-retinal filters, the sensitivity values for test wavelengths shorter than 480 nm were ignored. Inset, assessment of the number of X-chromosome-linked photopigments, showing the changes in equations for two test lights (540 and 630 nm) determined when the eye was alternately adapted to 540 and 630 nm lights; the scaling is logarithmic. Absence of a chromatic adaptation effect results in no shift in the equation (vertical broken line). Results are shown for three groups of subjects: howler monkeys (circles), normal human trichromats (open square) and human dichromats (solid square). For the human subjects, the plots are mean values (± 1 s.d.) for 14 normal trichromats and 27 dichromats (11 protanopes, 16 deuteranopes). **METHODS.** ERGs were recorded from sedated monkeys using contact-lens electrodes, and fibre electrodes for the human subjects. Stimuli flickering at 31.25 Hz were presented in maxwellian view. A flicker-photometric procedure was used in which the intensity of a flickering monochromatic light was adjusted in intensity until it produced an electrical response equivalent to that produced by the presentation of a flickering achromatic reference light^{25,26}. The two stimuli were interleaved in a temporal train. In the chromatic-adaptation experiment, the 540- and 630-nm adaptation lights were initially adjusted in intensity so that each produced a 0.5 log elevation in threshold for a 540-nm test light flickering at 31.25 Hz.

Old World primates. Polymerase chain reaction (PCR) amplification of exon 5 of the X-linked pigment genes from individuals having both long- and middle-wavelength pigment genes results in the formation of homoduplexes as well as heteroduplexes¹³⁻¹⁵. In individuals with only one type of pigment gene, only homoduplexes are formed. Heteroduplex analysis of amplified exon 5 of a male howler monkey (*A. seniculus*) was performed to see whether two different exon 5 sequences are present (Fig. 2a). Three bands are seen for the howler monkey: two heteroduplex bands and a homoduplex band. The presence of heteroduplexes indicates there are at least two different exon 5 sequences, and thus two different X-chromosome pigment genes. By comparison, a human dichromat has only a single exon 5 sequence and no heteroduplexes (Fig. 2a). Sequence ladders for exon 5 from two female and six male howler monkeys are shown in Fig. 2b. The sequences for all of the monkeys are very similar, but not identical: one female and one male have five sites with variations, whereas the remaining five males have four such sites. These locations of variation for all of these howler monkeys include the two locations that are correlated with the presence of two well-separated middle-to-long-wavelength cone pigments^{8,10,11}. To confirm that the combinations of nucleotides are also the same as in the Old World primates, exon 5 of the X-linked pigment genes amplified from genomic DNA from two male howler monkeys (males 3 and 4) was cloned into the plasmid vector pCR2.1 (Invitrogen), and individual clones were sequenced. The sequences fell into two classes with regard to the four locations of variation shown in Fig. 2b. In one class of clones,

TABLE 1 Nucleotide variations in exon 5 sequences for X-chromosome pigment genes

Nucleotide position	Howler monkeys	Old World monkeys	Other New World monkeys	Amino-acid substitution? (position)
1,295	G and A	G	G or A	No
1,314	A	T and A	A	No
1,317	A	C and T	G, C, A or T	Yes (275)
1,319	G and A	G and C	G	No
1,322	T and G	A and G	A or G	No
1,324	T and A	T and A	A or T	Yes (277)
1,329	G	T and G	A or G	Yes (279)
1,343	C	T and C	C	No
1,347	G and A	G and A	G or A	Yes (285)
1,410	C	T	C	Yes (302)

Nucleotide positions are shown where there were variations between howler monkeys, between the two pigment genes of two species of Old World monkey (*Cercopithecus diana* and *C. talapoin*)²², or between howler monkey and Old World monkeys. Exon 5 sequences were examined for the following New World monkeys: squirrel monkey (*Saimiri sciureus*)⁸, saddle-backed tamarin (*Saguinus fuscicollis*)⁸, common marmoset (*Callithrix jacchus*)²³, owl monkey (*Aotus trivirgatus*)²⁴, and spider monkey (*Ateles fusciceps*)²⁴.

the nucleotides at the four positions are: adenine at position 1,319; guanine at 1,322; thymine at 1,324; and guanine at 1,347. The other class of clones had guanine at position 1,319; thymine at 1,322; adenine at 1,324; and adenine at 1,347. The two classes of clones thus encode the same amino-acid combinations found in the human middle- and long-wavelength cone pigments. For male 3, three of five independent clones corresponded to the middle-wavelength pigment gene sequence, and two corresponded to the long-wavelength pigment gene sequence. For male 4, three of four independent clones correspond to long-wavelength pigment sequence, the other to middle-wavelength pigment gene sequence. The trichromatic colour vision of the howler monkey

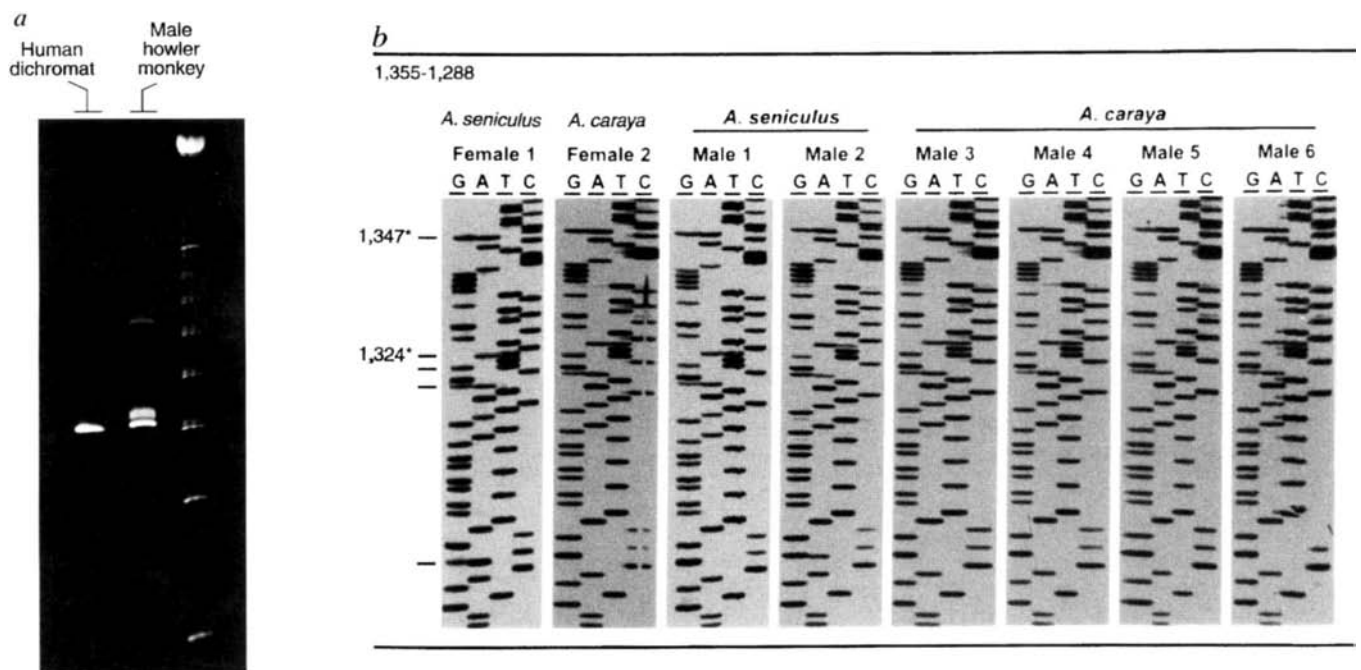


FIG. 2 a, Heteroduplex analysis of exon 5 of the X-linked pigment genes. Homoduplexes each contain both strands from a long-wavelength pigment gene, or both strands from a middle-wavelength pigment gene. The mismatches in the heteroduplexes alter their mobility relative to the homoduplexes. b, Sequencing autoradiograms spanning exon 5 nucleotides 1,355–1,288 for two female and six male howler monkeys. The numbering system is as ref. 17. Locations of variation are indicated by horizontal lines to the left; the two regions marked with asterisks specify

amino-acid differences associated with the presence of two types of cone pigments that are spectrally well separated in the middle to long wavelengths.

METHODS. DNA was isolated from blood and used in the PCR to amplify exon 5. The primers and PCR conditions are described elsewhere¹³. In a, amplified DNA was run on a 5% neutral polyacrylamide gel. The nucleotide sequences of PCR products were determined by cycle sequencing (Perkin Elmer).

implied by the electrophysiological results has a genetic basis like that of catarrhine primates. In these catarrhines, trichromacy is based on middle- and long-wavelength pigments with spectral peaks of about 530 and 562 nm (ref. 16). Predictions of sensitivity based on those pigments accounts well for the spectral sensitivity of the howler monkey (Fig. 1).

Trichromatic colour vision occurs in several other groups of animals, but primates are the only mammals to be trichromatic¹. The high sequence homologies of the X-chromosome pigment genes indicate that a gene duplication is required to convert the typical mammalian pattern (only a single X-chromosome pigment gene) to the multiple pigment genes of the Old World primates¹⁷. The discovery of widespread polymorphic colour vision among New World monkeys suggests that the evolution of routine trichromatic colour vision might involve an intermediate stage in which routine dichromacy is converted to polymorphic colour vision¹⁸. Alternatively, the colour-vision polymorphism of platyrrhine monkeys might be considered a degenerate form of catarrhine trichromacy¹⁹. Regarding howler monkeys, the difference between these two explanations depends on whether the trichromacy of howler monkeys arose from a polymorphic past, or whether it reflects the maintenance of an ancestral trichromacy. Two arguments support the former possibility. First, it is the more parsimonious way to explain the present picture: the phylogenetic relationships⁷ of the New World monkeys indicate that an ancestral trichromacy must have been independently lost in at least three lineages, whereas the evolution of trichromacy from a polymorphic base would have required change in only one line. Second, consider the 10 nucleotide positions in exon 5 that vary within two species of Old World monkey, within howler monkeys, or between Old World monkeys and howler monkeys. For each, the nucleotides and variants are given for five other species of New World monkeys in Table 1. For all comparisons there is no single case in which the sequence of the howler monkey is like the sequence of Old World monkeys but unlike that of other New World monkeys. For instance, there are positions (1,314 and 1,343) that vary in Old World monkeys but not in any of the New World monkeys; there is a location where all New World monkeys vary (1,295) but the Old World monkeys do not; and at a location where none of the species show variation (1,410), the howler monkey sequence is like that of other New World monkeys but unlike that of the Old World monkeys. Both arguments thus suggest that howler monkeys achieved their trichromacy as a result of evolutionary changes that occurred subsequent to the platyrrhine radiation, and that the capacity is constructed from a polymorphic past. Note that the only positions where all Old World and New World species show variation (1,324 and 1,347) are those yielding the amino-acid changes that are responsible for much of the difference in spectral positioning of middle- and long-wavelength cone pigments⁸. It has been suggested that the similarities at these locations across species represent convergent evolution²⁰.

It is not yet clear why howler monkeys are routinely trichromatic but other New World monkeys are not. The balanced polymorphic colour vision of the latter species means that groups of these monkeys include dichromatic and trichromatic individuals. It has been suggested that this arrangement is maintained because it provides a net group advantage in the detection of resources³. If that is the case, a detailed comparison of the visual ecology of the howler monkey with that of some of these other New World monkeys could prove informative. Alternatively, there might be a more mechanical explanation: perhaps the opsin gene duplication required for routine trichromacy is a low-probability event, achieved during the evolution of howler monkeys, but not as yet in other platyrrhine lines. In this case the constraint of a single X-chromosome pigment gene restricts a superior capacity for colour vision, trichromacy, to females. Female trichromats have higher fitness than dichromats, which in turn leads to stability in the polymorphism through heterozygous advantage²¹. □

Received 15 January; accepted 21 May 1996.

- Jacobs, G. H. *Biol. Rev.* **68**, 413–471 (1993).
- Jacobs, G. H. & Neitz, J. *Proc. natn. Acad. Sci. U.S.A.* **84**, 2545–2549 (1987).
- Tovee, M. J., Bowmaker, J. K. & Mollon, J. D. *Vision Res.* **32**, 867–878 (1992).
- Jacobs, G. H., Neitz, J. & Neitz, M. *Vision Res.* **33**, 269–274 (1993).
- Jacobs, G. H. & Deegan, J. F. *Invest. Ophthalmol. Vis. Sci.* **34**, 749 (1993).
- Jacobs, G. H. & Deegan, J. F. *Am. J. Primatol.* **33**, 217 (1994).
- Schneider, H. et al. *Molec. Phylogenet. Evol.* **2**, 225–242 (1993).
- Neitz, M., Neitz, J. & Jacobs, G. H. *Science* **252**, 971–974 (1991).
- Ibbotson, R. E., Hunt, D. M., Bowmaker, J. K. & Mollon, J. D. *Proc. R. Soc. Lond. B* **247**, 145–154 (1992).
- Chan, T., Lee, M. & Sakmar, T. P. *J. biol. Chem.* **267**, 9478–9480 (1992).
- Merbs, S. L. & Nathans, J. *Photochem. Photobiol.* **56**, 869–881 (1992).
- Williams, A. J., Hunt, D. M., Bowmaker, J. K. & Mollon, J. D. *EMBO J.* **11**, 2039–2045 (1992).
- Neitz, M. & Neitz, J. *Science* **267**, 1013–1016 (1995).
- Cotton, R. G. H. *Curr. Opin. Biotechnol.* **3**, 1013–1016 (1995).
- Nagamine, C. M., Chan, K. & Lau, Y.-F. *C. Am. J. hum. Genet.* **45**, 337–339 (1989).
- Bowmaker, J. K. in *From Pigments to Perception* (eds Valberg, A. & Lee, B. B.) 1–9 (Plenum, New York, 1991).
- Nathans, J., Thomas, D. & Hogness, D. S. *Science* **232**, 193–202 (1986).
- Bowmaker, J. K., Jacobs, G. H. & Mollon, J. D. *Proc. R. Soc. Lond. B* **231**, 383–390 (1987).
- Mollon, J. D. in *Evolution of the Eye and Visual System* (eds Cronly-Dillon, J. R. & Gregory, R. L.) 306–319 (CRC Press, Boca Raton, FL, 1991).
- Shyue, S.-K. et al. *Science* **269**, 1265–1267 (1994).
- Mollon, J. D., Bowmaker, J. K. & Jacobs, G. H. *Proc. R. Soc. Lond. B* **222**, 373–399 (1984).
- Dulai, K. S., Bowmaker, J. K., Mollon, J. D. & Hunt, D. M. *Vision Res.* **34**, 2483–2491 (1994).
- Hunt, D. M., Williams, A. J., Bowmaker, J. K. & Mollon, J. D. *Vision Res.* **33**, 147–154 (1993).
- Hagstrom, S. A. et al. *Invest. Ophthalmol. Vis. Sci.* **34**, 809 (1993).
- Neitz, J. & Jacobs, G. H. *J. opt. Soc. Am. A* **1**, 1175–1180 (1984).
- Jacobs, G. H., Neitz, J. & Krogh, K. J. *opt. Soc. Am. A* **13**, 641–648 (1996).

ACKNOWLEDGEMENTS. We thank C. Sweet, S. Evans, R. W. Cooper, T. Pope and S. Nikota for cooperation and assistance; C. Lauer for technical assistance; and J. A. Endler for comments. This work was supported by grants from the NSF, the NIH, and Research to Prevent Blindness.

CORRESPONDENCE and requests for materials should be addressed to G.H.J. (e-mail: jacobs@psych.ucsb.edu).

A PET study of the neural systems of stuttering

Peter T. Fox*, Roger J. Ingham†, Janis C. Ingham†, Traci B. Hirsch*, J. Hunter Downs*, Charles Martin*, Paul Jerabek*, Thomas Glass* & Jack L. Lancaster*

* Research Imaging Center and Department of Radiology, University of Texas Health Science Center at San Antonio, Texas 78284, USA

† Department of Speech and Hearing Sciences, University of California, Santa Barbara, California 93106, USA

THE cause of stuttering is unknown¹. Failure to develop left-hemispheric dominance for speech is a long-standing theory¹ although others implicate the motor system more broadly², often postulating hyperactivity of the right (language non-dominant) cerebral hemisphere³. As knowledge of motor circuitry has advanced⁴, theories of stuttering have become more anatomically specific, postulating hyperactivity of premotor cortex, either directly⁵ or through connectivity with the thalamus and basal ganglia⁶. Alternative theories target the auditory⁷ and speech production^{8,9} systems. By contrasting stuttering with fluent speech using positron emission tomography combined with chorus reading to induce fluency, we found support for each of these hypotheses. Stuttering induced widespread over-activations of the motor system in both cerebrum and cerebellum, with right cerebral dominance. Stuttered reading lacked left-lateralized activations of the auditory system, which are thought to support the self-monitoring of speech, and selectively deactivated a frontal-temporal system implicated in speech production. Induced fluency decreased or eliminated the overactivity in most motor areas, and largely reversed the auditory-system underactivations and the deactivation of the speech production system. Thus stuttering is a disorder affecting the multiple neural systems used for speaking.

Stuttering can often be transiently alleviated by a variety of interventions, collectively termed fluency inductions¹⁰. These are

either auditory stimulations (for example, chorus reading, masking and shadowing) or speech-pattern changes (such as singing, prolonged speech, rhythmic speech, whispering and shouting). This characteristic susceptibility of stuttering to fluency induction was used to investigate its neural circuitry.

Chorus paragraph reading was contrasted with solo paragraph reading in both normal controls and chronic developmental stutterers. Chorus reading completely eliminated stuttering in all ten stuttering subjects. In 30 chorus-reading scanning periods (ten subjects, three scans per subject), no subject was judged to have any stuttered intervals¹¹. In the same periods, the mean number of syllables spoken was 143.7 (range 121–173), not significantly different from the controls. During solo paragraph reading, all stutterers stuttered. The mean number of four-second intervals judged to be stuttered was 6.2 (range 1–10). The mean number of syllables spoken was 113 (range 82–154). In the non-stuttering controls, the mean number of syllables spoken during solo (mean 146.8, range 124–181) and chorus (mean 145.6, range 121–181) reading were not different. No controls had any stuttered intervals.

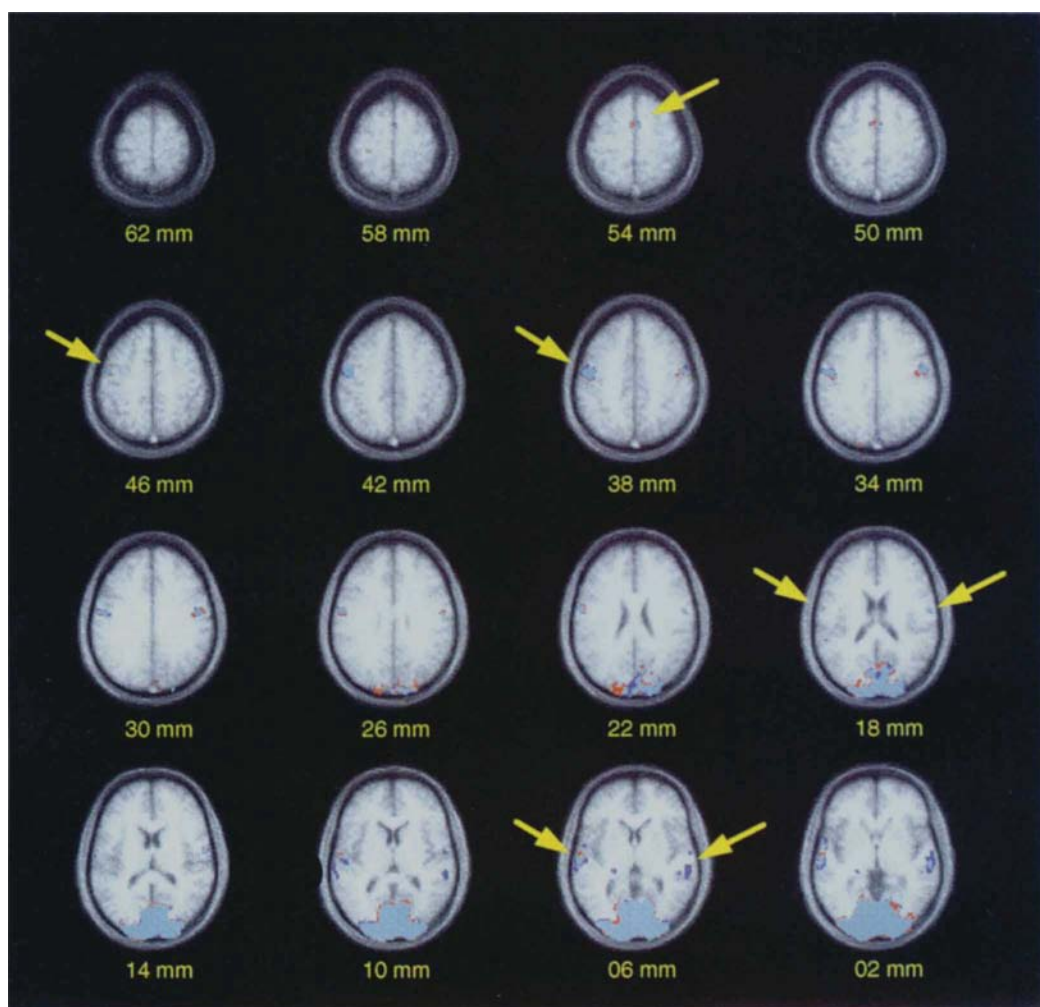
Positron-emission tomography (PET) was used to identify the

neural correlates of paragraph reading, stuttering and induced fluency. In normal controls, paragraph reading activated the primary motor cortex for mouth (M1, Brodmann area (BA) 4), supplementary motor area (SMA, medial BA6), inferior lateral premotor cortex (ILPrM, BA6/44), anterior–temporal extra-primary auditory cortex (A2, BA22/21), the visual system, and cerebellum (Fig. 1). Activations were predominantly left lateralized, with the exceptions of visual system and cerebellum, which were bilateral. These activations are similar to those of single-word reading, with two exceptions: SMA was less activated, and anterior temporal cortex (A2) was more activated, than previously reported^{12,13}. The greater activation of A2 is probably due to self-monitoring during continuous speech¹⁴. The decreased activation of SMA, despite the high complexity of the task and the much greater motor output, can be accounted for because SMA is principally responsible for motor-plan release, rather than motor programming *per se*^{15,16}. Specifically, we suggest that paragraph reading is a motor process (initiated only once) rather than a series of discrete motor acts (initiated per word).

Regional deactivations relative to rest, often unreported in the brain-mapping literature, were statistically significant and largely

FIG. 1 Brain activations for unaccompanied (solo) and accompanied (chorus) paragraph reading are illustrated in a novel, logical image format based on matrix application of logical operators to Z-score images. Coloured regions reflect areas of statistically significant change from rest. Activations common to both solo and chorus are coloured cyan and included SMA (arrow at 54 mm), SLPrM (arrow at 46 mm), M1 (arrow at 38 mm), ILPrM (arrows at 18 mm), left auditory cortex (left arrow at 06 mm) and visual cortex (posterior aspects of 22–02 mm, no arrows). Right A1 and A2 (right arrow at 18 mm) were far more active during chorus reading than solo (blue). Very few voxels were activated uniquely by unaccompanied reading (solo only; red). Planes are axial sections, labelled with the height (mm) relative to the bicommissural line. The left hemisphere is shown on the left.

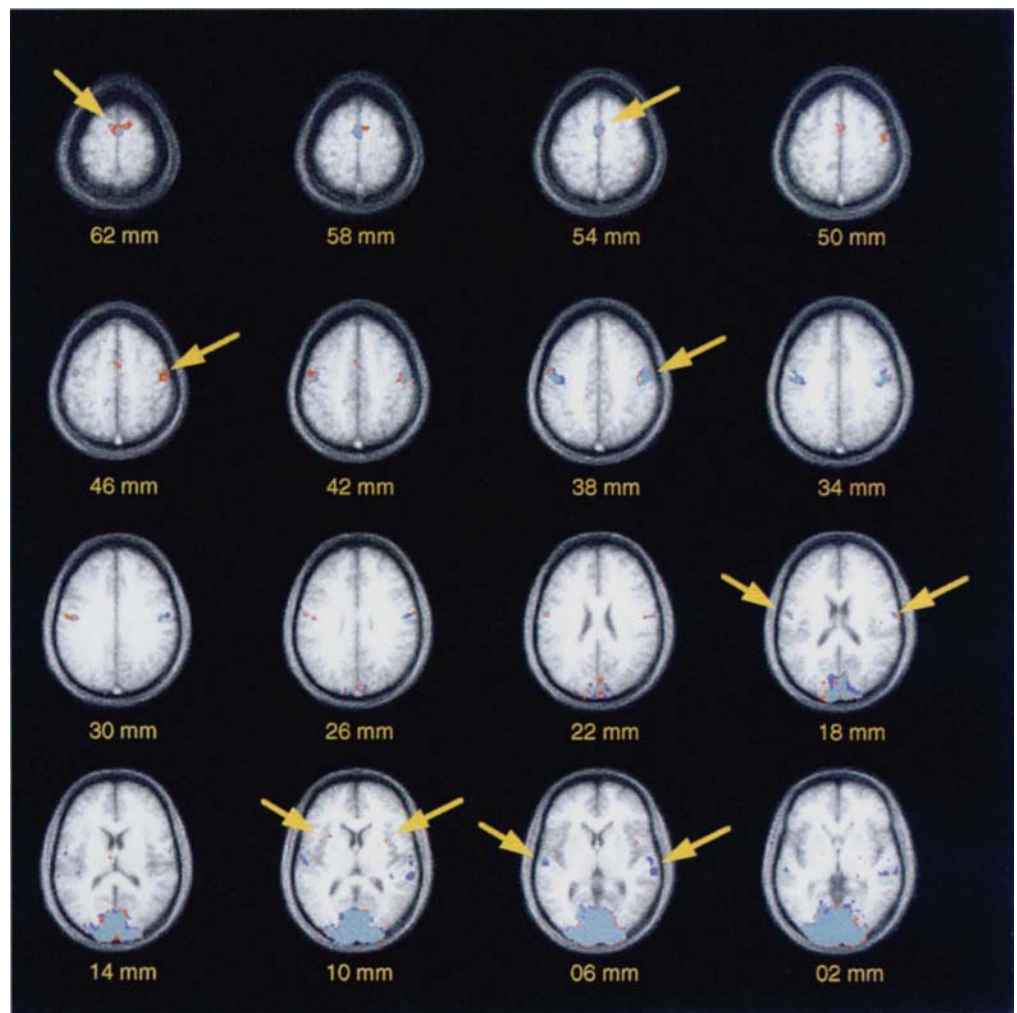
METHODS. Ten normally fluent, right-handed, adult (age 21–55 years, mean 32) men served as controls. Three conditions were studied: unaccompanied oral paragraph reading (solo), accompanied oral paragraph reading (chorus), and eyes-closed rest (rest). Task-induced changes in neural activity were detected as changes in the regional tissue uptake of ¹⁵O-water^{22,23}. Each condition was imaged three times in each subject. PET and magnetic resonance image (MRI) data were spatially normalized into the bicommissural coordinate space and averaged by condition and group^{24–26}. Grand-mean images for solo and chorus conditions were compared to rest, forming two group-mean subtraction images: solo (minus rest) and chorus (minus rest). After confirmation of omnibus significance by change-distribution analysis²⁶ and analysis of variance (ANOVA; condition ×



hemisphere × region, and also condition × hemisphere × region × group), subtraction images were converted to z-images (statistical parametric images of z-scores). z-images were converted to logical images by applying logical operators ('AND', 'NOT') to pairs of z-images, after eliminating all voxels below a z-score value of 1.96 ($P < 0.01$). In this way, each activated voxel was placed in one of three mutually exclusive categories: solo and chorus; solo not chorus; or chorus not solo.

FIG. 2 Logical images (see Fig. 1, Methods) for unaccompanied (solo) and accompanied (chorus) paragraph reading in stutterers. Stuttering (solo not chorus, red) was uniquely associated with over-activity of SMA (arrows at 62 mm and 54 mm), SLPrM (arrow at 46 mm) and insula (arrows at 10 mm), among others. In both conditions (solo and chorus, cyan), primary motor cortex for mouth was more activated on the right (arrow at 38 mm), contra-lateral to the predominant activation in controls. Induced fluency (chorus not solo, blue) greatly reduced the intensity of the activations in the SMA (arrow at 62 mm) and superior premotor cortex (arrow at 46 mm), but did not normalize the laterality of M1 (arrow at 38 mm).

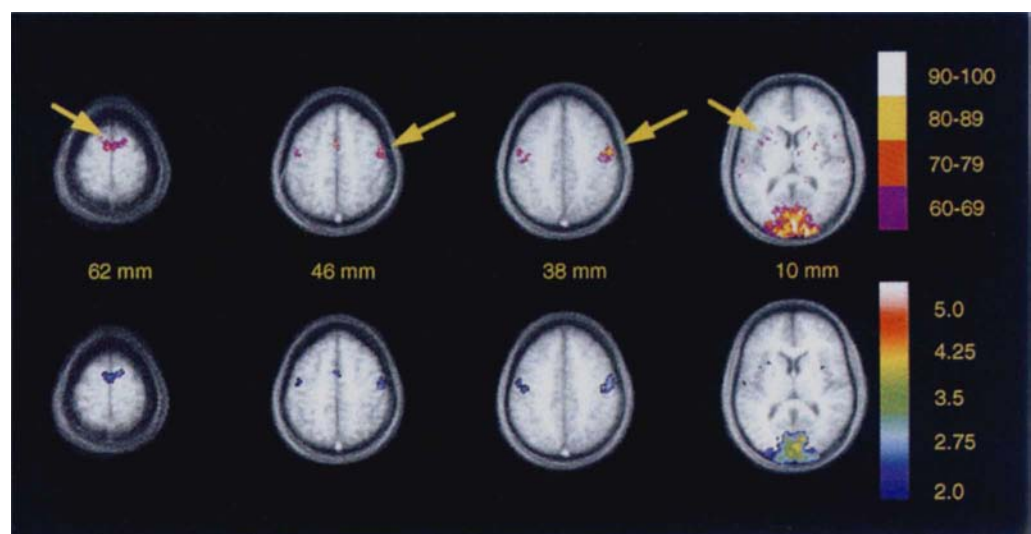
METHODS. Ten right-handed, adult (age 21–46 years, mean 32) men with chronic, developmental stuttering served as subjects. As in the controls, three conditions were studied: solo, chorus, and rest. Paragraphs (identical in solo and chorus) were presented on a video monitor 14 inches from the eyes. In chorus, a tape recording of a non-stutterer reading the same passage was played through an earphone in the left ear at a rate chosen by the stutterer to be most comfortable²⁷. For each stutterer, a randomly selected control heard (and read) the passage at the same rate. For both patients and controls, speech was recorded during each of the nine, 40-s, PET-acquisition periods. These 40-s recordings were divided into ten 4-s epochs, each of which was scored for the presence of stuttering and the number of syllables uttered. Epochs were scored by R.J.I. and independently by a clinician at another university^{11,28}. Across subjects, interval-by-interval agreement regarding the presence or absence of stuttering ranged



from 85.0% to 100%, with 100% agreement that no stutterer stuttered during chorus. Total percentage agreement for syllables ranged from 97.9% to 99.7%. Conditions were counterbalanced to control for order and stuttering-adaptation effects²⁹; neither was present.

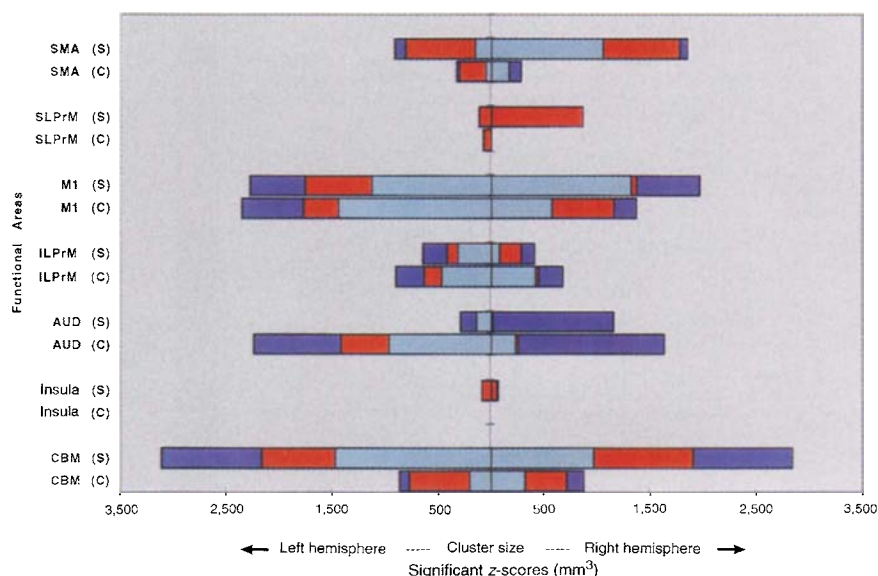
FIG. 3 The intersubject consistency of regional brain activations associated with stuttering is illustrated in a format termed penetrance images. Planes are axial sections, labelled with the height (mm) relative to the bicommissural line. The left hemisphere is shown on the left. Top row, penetrance values (ranging from 60% (purple) to 90% (white)) are displayed. In the right column, the mean intensity of the activation is displayed as z-scores (range, 2–5) (see Fig. 1, Methods). Penetrance images and z-images are overlaid on the stutterers' group-mean MRI. The stuttering-specific overactivations of SMA (arrow at 62 mm), SLPrM (arrow at 46 mm), right M1 (arrow at 38 mm) and insula (arrow at 10 mm) achieved peak penetrance values of 80–100%.

METHODS. Penetrance images were created from individual z-images (Fig. 1, Methods) of the stutterers' solo (minus rest) condition. z-images were converted to binary images (each voxel valued either 0 or 1), based on a z-threshold of $P < 0.1$. The binary images were summed across the ten



subjects and multiplied by 10, to convert the values to percentages. In non-activated areas, a mean penetrance of 10% (range 0–30%) was observed, as predicted by binomial sampling theory.

Fig. 4 The volume (mm^3) of the regional activations observed in each brain region for each subject group and logical contrast (see Figs 1 and 2) is illustrated as a stacked bar graph. For each brain region, stutterers (S) are shown in the upper bar; controls (C) are shown in the lower bar. The volume activated by both conditions (solo and chorus) is coloured cyan; the volume activated only by unaccompanied reading (solo not chorus) is red; the volume activated uniquely by accompanied reading (chorus not solo) is blue. Extents are plotted to be cumulative, not overlapping. Thus total activation extent during solo is cyan plus red (that is, solo and chorus, plus solo not chorus.) Total activation extent during chorus is cyan plus blue (that is, solo and chorus, plus chorus not solo). Note that, overall, motor activations (SMA, SLPrM, insula and cerebellum) are much more extensive in stutterers than in controls. Conversely, auditory activations are much more extensive in controls than in stutterers. The overactivation of SMA, SLPrM and insula seen during stuttering was greatly reduced during induced fluency (chorus reading). METHODS. The logical images (solo and chorus; solo not chorus; chorus not solo) displayed in Figs 1 and 2 were searched for local maxima³⁰. By using the local maximum as a search kernel, the number of contiguous pixels with a z-value above 1.96 and lying within a 5-mm radius was determined. These clusters of activated pixels were identified by Brodmann area and gyrus, based on the designations of



Talairach²⁵. For each functional area, the total extent of all activated clusters was summed.

consistent across tasks and groups. In solo paragraph reading in the controls, deactivations were observed in the motor system (insula, putamen and premotor cortex (BA9)), the memory system (left hippocampal gyrus (BA34) and right lateral prefrontal cortex (BA47)) and the attentional system (multiple sites throughout the cingulate gyrus (BA10, 32, 24, 31 and 7)).

During chorus reading of the same paragraphs, controls again activated M1, SMA, ILPrM and A2, predominately on the left, although the right auditory cortex (A1, BA41; A2 BA22) was also activated, attributable to the left-ear auditory stimulation used for the chorus-reading procedure (Fig. 2, Methods). Deactivations in chorus reading were largely identical to those of solo reading.

In stuttered reading, neural-system activations differed markedly from those of the controls (Fig. 2). These differences were largely in the motor system, as follows. M1 was similar to the controls in extent and intensity, but was right lateralized. SMA was far more active than in the controls, with two large, discrete foci: one focus was similar in location to the SMA in controls, although it was larger and more intense; a second focus was very extensive and more than 1 cm superior to the normally located focus. Superior lateral premotor cortex (SLPrM; lateral BA6), minimally activated in the controls, was strongly activated and strongly right lateralized. Cerebellar activations were of more than double the extent of the controls, in either condition. Motor-system activations not seen in the controls included the insula (bilaterally), claustrum (left), lateral thalamus (left) and globus pallidus (left). ILPrM was the only motor region that activated in a 'normal' manner, being left lateralized and similar in intensity and extent to the controls. Thus stuttering was characterized by extensive hyperactivity of the motor system, with right lateralization of primary and extraprimory motor cortices. Left superior temporal activations, observed in the controls and attributed to self-monitoring, were virtually absent during stuttering.

Deactivations during stuttering were also distinctive. Not only did left superior temporal cortex fail to activate (above), but left posterior temporal cortex (BA22) showed significant deactivations not seen in the controls. Furthermore, the left inferior frontal cortex (BA47), implicated in verbal comprehension^{12,17} and verbal fluency^{18,19}, was focally deactivated. The deactivation of left posterior temporal and left inferior frontal cortex, taken

together with the failure to activate normally left superior temporal cortex, strongly implicates a frontal-temporal system for verbal fluency²⁰.

Regional activations and deactivations observed during stuttering were remarkably consistent. Stuttering-specific activations in SMA, SLPrM and insula, for example, were present in 70–100% of subjects (Fig. 3). This consistency is remarkable when considering that stuttering is often characterized as being highly variable¹. Among stutterers, severity ranges from barely detectable to severely disabling. (Our subjects ranged in degree of stuttering from mild to severe.) Even within a single subject, the frequency of stuttering often varies widely in the course of a day.

Induced fluency markedly reduced the abnormalities seen in stuttering (Fig. 4). Motor-system overactivations were reduced (in the SMA and cerebellum) or eliminated (in the SLPrM, insula, claustrum and globus pallidus). Deactivations of left inferior frontal and left posterior temporal cortex were eliminated, and lack of activation in left superior temporal cortex was substantially reduced. Chorus reading, however, did not normalize the hemispheric dominance of M1 or SMA.

In summary, we have mapped the neural systems of normal continuous speech, both activations and deactivations. The neural systems of stuttering have been isolated and include: (1) diffuse overactivity of the cerebral and cerebellar motor systems; (2) right dominance of the cerebral motor system; (3) lack of normal 'self-monitoring' activations of left, anterior, superior temporal phonological circuits; and (4) deactivation of a 'verbal fluency' circuit between left frontal (BA47) and left temporal (BA22) cortex²⁰. Reminiscent of the parable of the blind men and the elephant²¹, current theories of stuttering each emphasize one or another individual component of what we believe to be a dysfunctional system or systems. Our results lend support to several prominent theories of stuttering, and strongly indicate the need for a unifying theory of sufficient scope to accommodate the full complexity of the observed actions and interactions of the neural systems. □

Received 31 May 1995; accepted 10 May 1996.

1. Bloodstein, O. *A Handbook on Stuttering* (Singular Group, San Diego, 1995).
2. Zimmermann, G. J. *Speech Hear. Res.* **23**, 122–136 (1980).
3. Yeudall, L. T. *Semin. Speech Lang.* **6**, 197–231 (1985).
4. Alexander, G. E., DeLong, M. R. & Strick, P. L. *Rev. Neurosci.* **9**, 357–381 (1986).

5. Webster, W. G. in *Neuropsychology of Stuttering* (ed. Boberg, E.) 73–127 (University of Alberta, Edmonton, 1993).
6. Caruso, A. J. in *Speech Motor Control and Stuttering* (eds Peters, H. F. M., Hulstijn, W. & Starkweather, C. W.) 101–116 (Excerpta Medica, Amsterdam, 1991).
7. Stromsta, C. *Elements of Stuttering* (Oshtemo Michigan, Atsmons, 1986).
8. Perkins, W. H., Kent, R. & Curlee, R. D. *J. Speech Hear. Res.* **34**, 734–752 (1991).
9. Wingate, M. *The Structure of Stuttering: A Psycholinguistic Analysis* (Springer, New York, 1988).
10. Ingham, R. J. *Stuttering and Behavior Therapy: Current Status and Experimental Foundations* (College Hill, San Diego, 1984).
11. Ingham, R. J., Cordes, A. K. & Finn, P. J. *Speech Hear. Res.* **36**, 1168–1176 (1993).
12. Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M. & Raichle, M. E. *J. cogn. Neurosci.* **1**, 153–170 (1989).
13. Pardo, J. V. & Fox, P. T. *Hum. Brain Mapp.* **1**, 57–68 (1993).
14. Fairbanks, G. *J. Speech Hear. Disord.* **19**, 133–139 (1954).
15. Fox, P. T., Fox, J. M., Raichle, M. E. & Burde, R. M. *J. Neurophysiol.* **52**, 348–368 (1985).
16. Grafton, S. T. *Ann. Neurol.* **36**, 3–4 (1994).
17. Posner, M. I., Petersen, S. E., Fox, P. T. & Raichle, M. E. *Science* **240**, 1627–1631 (1988).
18. Frith, C. D., Friston, K. J., Liddle, P. F. & Frackowiak, R. S. J. *Neuropsychologia* **29**, 1137–1148 (1991).
19. Wise, R. J. et al. *Brain* **114**, 1803–1817 (1991).
20. Friston, K. J., Frith, C. D., Liddle, P. F. & Frackowiak, R. S. J. *Proc. R. Soc. Lond. B* **244**, 101–106 (1991).
21. Zimmermann, G. in *Nature and Treatment of Stuttering: New Directions* (eds Curlee, R. W. & Perkins, W. H.) 131–147 (College Hill, San Diego, 1984).
22. Fox, P. T., Mintun, M. A., Raichle, M. E. & Herscovitch, P. J. *cereb. Blood Flow Metab.* **4**, 329–333 (1984).
23. Fox, P. T. & Mintun, M. A. *J. nuclear Med.* **30**, 141–149 (1989).
24. Fox, P. T., Perlmutter, J. S. & Raichle, M. E. *J. Comput. assist. Tomogr.* **9**, 141–153 (1985).
25. Talairach, J. & Tournoux, P. *Co-planar Stereotaxic Atlas of the Human Brain* (Springer, Thieme, 1988).
26. Fox, P. T., Mintun, M. A., Reiman, E. M. & Raichle, M. E. *J. cereb. Blood Flow Metab.* **8**, 642–653 (1988).
27. Ingham, R. J. & Packman, A. J. *Speech Hear. Res.* **22**, 784–793 (1979).
28. Cordes, A. K. & Ingham, R. J. *J. Speech Hear. Res.* **38**, 33–41 (1995).
29. Van Riper, C. & Hull, C. J. in *Stuttering in Children and Adults* (eds Johnson, W. & Leutenegger, R. R.) 199–206 (Univ. Minnesota Press, Minneapolis, 1955).
30. Mintun, M. A., Fox, P. T. & Raichle, M. E. *J. cereb. Blood Flow Metab.* **9**, 96–103 (1989).

ACKNOWLEDGEMENTS. We thank B. Alkek and K. Myers for recruiting and screening stutterers; R. Moglia, B. Heyl, R. Evans and F. Zamarripa for technical assistance; B. Alkek for quantifying speech performance; L. Parsons, H. Mayberg, M. Liotti and M. Woldorff for criticism of this manuscript and the new image formats (Penetration and Logical Images).

CORRESPONDENCE and requests for materials should be addressed to P.T.F. (e-mail: fox@uthscsa.edu). Data reported here are available through the BrainMap Database (bmap.admin@uthscsa.edu).

referred to as the anterior and posterior compartments. In leg imaginal discs, Hh protein in the posterior compartment induces the expression of high levels of Dpp dorsally in a stripe of cells immediately anterior to the compartment boundary (Fig. 1*f, h*). In wing discs, the Hh-induced stripe of Dpp extends the length of the compartment boundary¹. Hh also induces Wg expression and low

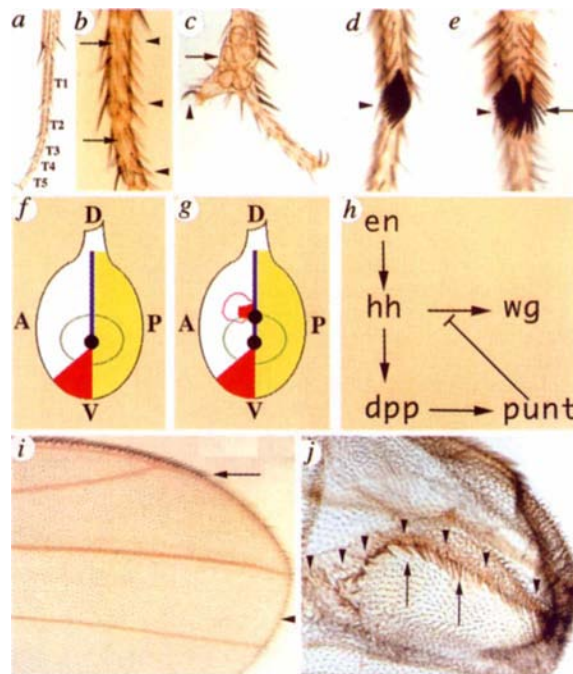


FIG. 1 *punt* clones cause pattern reorganization in adult legs and wings. Anterior view of adult legs with distal down, and proximal up (a–e). a, Distal tibia, five tarsal segments and claw of a wild-type second leg. b, Ventral or lateral *punt* clones marked by the gene yellow (*y*) (arrows and arrowheads) do not cause abnormalities in adult legs. Ventral bristles (arrowheads) are short and stout, whereas dorsal bristles (far left) are long and tapered. c, Dorsal *punt* clones can cause bifurcations. In the clone, bristles have the morphology of ventral or lateral bristles, even though they are in a dorsal position. d, Wild-type male first leg to indicate the sex comb (arrowhead) in an anterior ventral-lateral position. e, Dorsal *punt* clones (*y* bristles) induce sex-comb duplications (arrow). The endogenous sex comb is on the left (arrowhead). *punt* clones can reorganize wild-type tissue. Some of the bristles in the ectopic sex comb are *y*⁺, and in c most of the bristles on the supernumerary outgrowth are *y*⁺ (arrowhead below the claw). f, The tarsal segments originate from the end-knob region of the imaginal disc (green circle) and the most distal structure, the claw, originates from the centre of the disc (indicated by a black circle). In response to *hh* protein (yellow) from the posterior compartment, anterior compartment cells express Dpp (blue) at high levels dorsally, and Wg (red) ventrally¹. Although high levels of *dpp* and *wg* do not overlap, their juxtaposition at the centre of the leg disc (black circle) promotes distal outgrowth^{7,26,27}. g, A *punt* clone occurring just anterior to the A/P boundary induces ectopic *wg* expression and, in some cases, the formation of a supernumerary outgrowth. h, Hh expression is activated by the *en* selector gene in the posterior compartment. In dorsal distal regions we suggest that *dpp* acts through *punt* to repress Hh-induced *wg* expression along the anterior side of the compartment boundary. i, A wild-type wing with anterior to the top and posterior down. The A/P compartment boundary occurs between veins 3 and 4 (arrowhead) and the margin bristles on the periphery of the wing are visible. Stout mechanosensory bristles are present in the anterior compartment (arrow). j, A wing, oriented as in i, which contains a *punt* clone marked by *y* (arrowheads) in the middle of the wing in a position that is close to the A/P boundary. The ectopic margin contains anterior stout mechanosensory bristles (arrows), suggesting that it is in the anterior compartment. *punt* clones that arise on the margin in positions that are far from the A/P boundary do not have severe defects. METHODS. Clones were induced as for Fig. 2, except that males were of the genotype *yw P{ry[+] hsp70 - flp}; P{ry[+] hsp70 : neo FRT}82B punt⁶⁶²/CXD* and larvae were allowed to reach adulthood.

Decapentaplegic restricts the domain of wingless during *Drosophila* limb patterning

Andrea Penton & F. Michael Hoffmann

McArdle Laboratory for Cancer Research and Laboratory of Genetics, University of Wisconsin Medical School, Madison, Wisconsin 53706, USA

SIGNALLING proteins in the BMP-decapentaplegic (*dpp*), WNT-wingless (*wg*) and Shh-hedgehog (*hh*) families have been implicated in limb and appendage development in both invertebrates and vertebrates^{1–3}. In *Drosophila*, *dpp* protein (Dpp) induces distal outgrowth and patterning of legs and wings, but the molecular responses to Dpp are not well characterized^{4–9}. Analysis of clones mutant for the Dpp receptors encoded by *punt* or *thickveins* (*tkv*) reveals that repression of *wg* expression is one critical function of Dpp signalling in leg and wing discs. Distal clones that lie on the anterior edge of the anterior–posterior compartment boundary ectopically express *wg* and cause pattern abnormalities, suggesting that Dpp represses Hh activation of *wg* in the distal primordia of the leg and wing. By repressing *wg* expression in the leg, Dpp signalling limits the region that responds to high levels of Wg and Dpp to the site of distal outgrowth. Such negative regulatory feedback loops between signalling molecules are likely to be critical for limb patterning in other species.

During *Drosophila* development, undifferentiated epithelial cells in the imaginal discs proliferate and undergo patterning events that result in the formation of the adult appendages at metamorphosis. These cells belong to two distinct cell lineages,

levels of Dpp in the ventral anterior quadrant of the leg disc (Fig. 1f, h). In the leg disc, distal outgrowth occurs from the centre of the disc where Dpp and Wg protein are juxtaposed; *dpp* is necessary and sufficient to organize patterning in leg and wing imaginal discs^{4,9}.

Cells respond to Dpp through a heteromeric complex consisting of a type I receptor, either thickveins (Tkv) or saxophone (Sax), and a type II receptor, *punt*¹⁰⁻¹⁴. To investigate the role of Dpp signalling in disc pattern formation, clones of cells mutant for *punt* or *tkv* were induced in 2nd-instar larvae using flip-mediated somatic recombination^{15,16}. The *punt* allele (*punt*^{P62}) used here is not a null allele¹⁴, but yielded a higher frequency of clones and larger clones than *punt*¹³⁵ or null alleles of *tkv* (A.P. and F.M.H., manuscript in preparation). Thus the effects observed in the *punt*-mutant cells are probably due to a quantitative reduction in Dpp signalling in the responding cells rather than to a complete absence of Dpp signalling.

In the leg, *punt* clones that arose in ventral regions or that were proximal to the distal tibia had no effect on the bristle pattern (Fig. 1b and Table 1). However, *punt* clones located on the extreme dorsal surface of distal leg segments were associated with a loss of dorsal structures, duplication of ventral structures, and dorsal bifurcations (Fig. 1c, e and Table 1). The duplicated structures frequently contained wild-type bristles, indicating that dorsal *punt* clones could reorganize the surrounding wild-type tissue (Fig. 1c, e). Adult wings that contained *punt* clones had defects, including reduced wing size, ectopic vein formation, blisters and, most strikingly, ectopic wing margins in the middle of the blade (Fig. 1j).

Clones mutant for *punt* (no Myc staining) were recovered in both the anterior and posterior compartments of the leg disc (Fig. 2 and Table 1). Although the pattern of Wg in the ventral-anterior sector was not altered in *punt* clones, clones in the dorsal, distal primordia of the leg disc inappropriately expressed *wg* on the anterior side of the compartment boundary (Fig. 2a-h). Clones that were in the posterior compartment, that did not abut the anterior-posterior (A/P) compartment boundary, or that were proximal of the area fated to become the tibia, did not ectopically express *wg* (Fig. 2a-h and Table 1). *tkv*-null clones yielded similar results to *punt* clones, indicating that our observations were not due to an inherent peculiarity of the *punt*^{P62} allele (Fig. 2i-l and Table 1).

In the wing disc, *wg* is initially expressed in the entire ventral half of the disc and, during the third instar, resolves into a prominent stripe of expression along the presumptive wing margin, a ring of expression in the presumptive hinge regions (Fig. 3f), and a patch of expression in the notum primordia¹⁷. These expression patterns of *wg* are not known to be regulated by Hh signalling. In wing discs, *punt* clones ectopically expressed *wg* along the anterior side of the compartment boundary in either the dorsal or ventral portion of the pouch region (Fig. 3a-d and Table 1). In almost all cases, *punt* clones did not ectopically express *wg* if they did not abut the A/P compartment boundary, arose in the

TABLE 1 Number of clones scored in specific regions of leg discs, wing discs and adult legs

	Leg disc clones		Number with detectable <i>wg</i> expression		Wing disc clones	Number with detectable <i>wg</i> expression
	<i>punt</i>	<i>tkv</i>	<i>punt</i>	<i>tkv</i>	<i>punt</i>	<i>punt</i>
Number of clones	111	48	18	6	119	9
Non-A/P boundary clones:						
Anterior	52	28	0	0	57	0
Posterior	17	4	0	0	30	0
A/P boundary clones:						
Anterior dorsal proximal	5	0	0	0	13	0
Anterior dorsal distal	19	6	18	6	6	4
Anterior ventral	7	4	*	*	5	4
Posterior	11	6	0	0	8	1
Adult legs						
	Number	Number with defects				
Number of legs	37	19				
Proximal <i>punt</i> clones	177	1				
Ventral or lateral distal <i>punt</i> clones	120	1				
Dorsal distal <i>punt</i> clones	23	17				

The *punt* and *tkv* disc clones were generated and marked as for Fig. 2, and adult *punt* clones as for Fig. 1. The number of leg discs included in this analysis was 41 and 25 for *punt* and *tkv*, respectively. We examined 40 wing discs for *punt* clones. Discs that did not possess clones are not included here. The high frequency of clones and twospots (Fig. 2) made it difficult to determine in some cases whether a clone was actually two clones. In the cuticle, if *y* bristles were interrupted by *y*⁻ bristles, clones were scored as two. *tkv* clones were not recovered in the wing pouch and at lower frequencies than *punt* clones in leg discs. The frequency and size of *punt* clones in the cuticle did not differ significantly from wild-type control clones. The smallest *punt* clones scored in the adult consisted of one *y* bristle and the largest was 36 *y* bristles. In the adult, small (one or two *y* bristles) dorsal distal clones usually did not have a phenotype, whereas clones that were larger almost always possessed abnormalities.

* *punt* and *tkv* clones that occurred in anterior ventral positions in the leg disc did not alter the normal pattern of *wg* expression.

posterior compartment, or occurred proximally (Fig. 3a-h and Table 1).

Normal pattern formation requires a mechanism to limit the regions in which *dpp* and *wg* expression overlap (Fig. 1f, g). Hh induces *wg* expression in the embryo and in the ventral anterior region of the leg disc^{1,18}. Hh also induces *dpp* expressions along the anterior side of the compartment boundary in leg and wing discs. Consequently, the potential exists for extensive overlap in *wg* and *dpp* expression along the A/P compartment boundary. However, this overlap would be detrimental to patterning in the leg disc because regions in which high levels of Dpp and Wg are expressed in close proximity induce distal outgrowth (Fig. 1f, g), and experimentally induced ectopic sites of Wg-Dpp intersections cause leg duplications^{1,6,7}. By repressing expression of *wg* dorsally, Dpp limits leg outgrowth to the centre of the disc.

In addition to limiting the position of distal outgrowth, Dpp is necessary to limit patterning events controlled by Wg. Ectopic *wg* expression causes the formation of supernumerary ventral structures in legs and margin bristles in wings¹⁹⁻²². We propose that Dpp restricts the position of ventral structures by repressing *wg* expression dorsally in leg discs. In the wing blade, we posit that Dpp limits *wg* expression to Hh-independent regions at the dorsal-ventral boundary and thus confines the formation of margin bristles to the circumference of the wing.

Repression of *wg* expression represents a new role for *dpp* in disc development and raises the possibility that similar negative regulatory feedback loops occur during the vertebrate limb growth. Shh, BMPs, Wnts and fibroblast growth factors influence patterning of the vertebrate limb, a process that must involve integration of multiple signals to specify particular cell fates^{2,3}. Positive regulatory interactions have been established for some of these factors²³⁻²⁵. We propose that negative regulatory interactions are likely to be equally important in limiting which regions

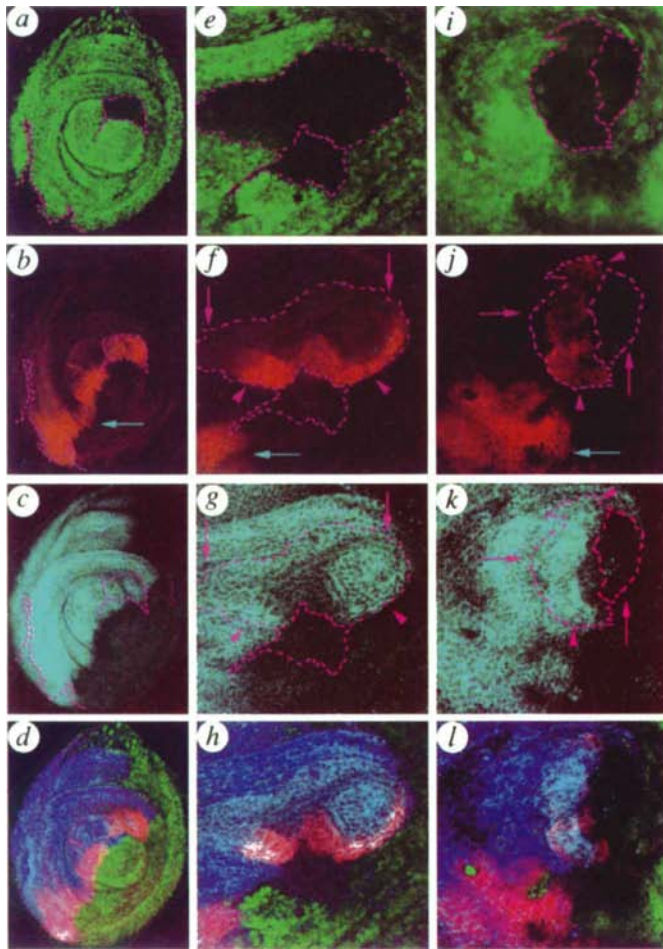


FIG. 2 Distal *punt* or *tkv* clones that lie anterior to the compartment boundary ectopically express *wg* in leg discs. Confocal images of discs that were triple-labelled to indicate *punt* or *tkv* clones (loss of Myc staining in green; *a*, *e* and *i*), *wg-lacZ* expression (anti- β -galactosidase in red; *b*, *f* and *j*), and the anterior compartment (cubitus interruptus, *ci*) expression in blue; *c*, *g* and *k*). Three independent discs are indicated in *a-d*, *e-h* and *i-l* (with anterior to the left, posterior to the right, dorsal up and ventral down). *a*, Two *punt* clones (pink outlines) are indicated. The dorsal clone (right) is located distally near the centre of the disc in a region that will form the distal tibia and tarsal segments, and abuts the compartment boundary from the anterior side (*c*). *wg* is normally expressed in the anterior ventral portion of the disc (blue arrow in *b*, *f* and *j*)¹⁷ and is ectopically expressed in the dorsal clone (*b*). The anterior ventral clone (left, *a*) does not misexpress *wg* and does not appear to disrupt the endogenous pattern of *wg* expression (*b*). *e*, Two *punt* clones based on the presence of two twin spots from the reciprocal recombination class are indicated (intense staining). The disc is distorted and appears to be forming another organizing centre. Only those cells that are most distal and lie next to the compartment boundary ectopically express *wg* (arrowheads in *f* and *g*). Cells that are more proximal or far from the compartment boundary fail to express *wg* (pink arrows in *f* and *g*). The posterior clone (absence of *ci* staining; *g*) fails to express *wg*. *i*, Two *tkv* clones. Only the anterior clone expresses *wg* (pink arrowheads in *j* and *k*). Anterior cells that are far from the compartment boundary or posterior cells (arrows in *j* and *k*) fail to express *wg*. *d*, *h*, *l*, A merge of *a-c*, *e-g* and *i-k*, respectively. The overlap of *wg-lacZ* and *ci* staining is shown in pink and indicates that both normal and ectopic *wg* expression are confined to the anterior compartment.

METHODS. Males of the genotype *w, wg-lacZ/In(2LR)Gla Bc, P{ry[+]hsp70: neo FRT}82B punt^{tr62}/TM6B* were crossed to females of the genotype *yw P{ry[+]hsp70-flip}; P{ry[+]hsp70: neo FRT}82B P{w[+]m}hsp70 mycPiM}87E Sb P{y[+], ry[+]}96E* to obtain *punt* clones. To obtain *tkv* clones, males of the genotype *w, tkv⁵ wg-lacZ P{ry[+]hsp70: neo FRT}40A/In(2LR)Gla Bc* were crossed to females of the genotype *yw P{ry[+]hsp70-flip}; P{w[+]m}hsp70 mycPiM}21C P{PiM}36F P{ry[+]hsp70: neo FRT}40A*. The staining was done by either a two- or three-step method^{28,29} using a fluorescein conjugate to visualize anti-Myc antibody, a rhodamine conjugate for the anti-*ci* antibody, and a Cy-5 conjugate for the anti- β -gal antibody.

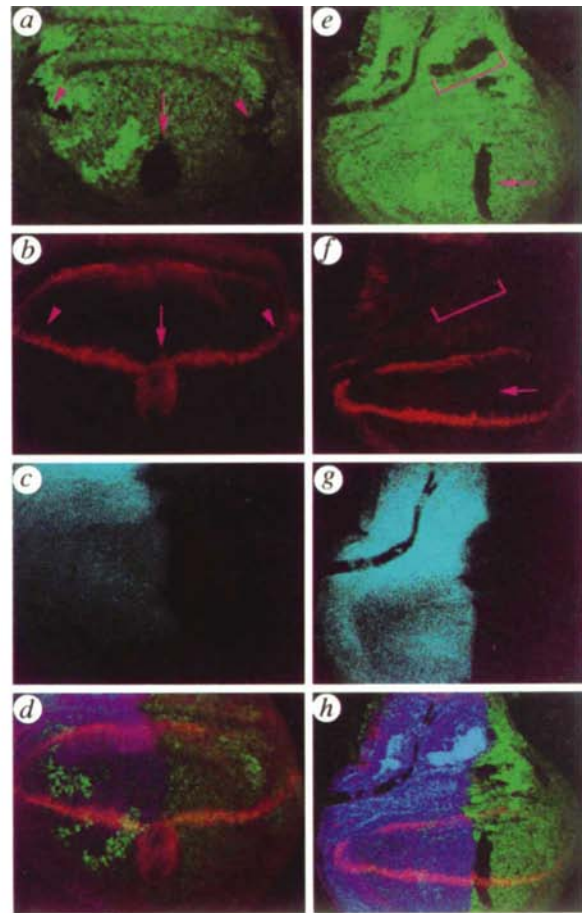


FIG. 3 *punt* clones that are anterior to the A/P compartment boundary ectopically express *wg* in the wing pouch. Discs were triple-labelled as for Fig. 2. Two independent discs are shown in *a-d* and *e-h*. Myc staining is in green (*a*, *e*), *wg-lacZ* staining in red (*b*, *f*), and *ci* staining in blue (*c*, *g*). Discs are oriented as in Fig. 2. Ectopic *wg* expression is detected in the ventrally situated central *punt* clone (arrows in *a* and *b*) but not in two outer clones (arrowheads in *a* and *b*). The central clone is located anterior to the compartment boundary (*c* and *d*). The intensely staining patch of cells anterior to this clone is generated by the reciprocal recombination class or twin spot. *e-h*, *punt* clones that are on the posterior side of the compartment boundary (arrows in *e* and *f*) fail to express *wg* ectopically. Inductions of clones and antibody stainings were done as in Fig. 2.

express specific signalling molecules, thereby restricting the position of specific limb structures. □

Received 18 April; accepted 30 May 1996.

- Basler, K. & Struhl, G. *Nature* **368**, 208–214 (1994).
- Kingsley, D. M. *Genes Dev.* **8**, 133–146 (1994).
- Tickle, C. *Curr. Opin. Genet. Dev.* **5**, 478–484 (1995).
- Spencer, F. A., Hoffmann, F. M. & Gelbart, W. M. *Cell* **28**, 451–461 (1982).
- Masucci, J. D., Miltenberger, R. J. & Hoffmann, F. M. *Genes Dev.* **4**, 2011–2023 (1990).
- Capdevila, J. & Guerrero, I. *EMBO J.* **13**, 4459–4468 (1994).
- Diaz-Benjumea, F. J., Cohen, B. & Cohen, S. M. *Nature* **372**, 175–179 (1994).
- Zecca, M., Basler, K. & Struhl, G. *Development* **121**, 2265–2278 (1995).
- Ingham, P. W. & Fietz, M. J. *Curr. Biol.* **5**, 432–440 (1995).
- Brummel, T. J. et al. *Cell* **78**, 251–262 (1994).
- Nellen, D., Affolter, M. & Basler, K. *Cell* **78**, 225–237 (1994).
- Penton, A. et al. *Cell* **78**, 239–250 (1994).
- Ruberte, E., Marty, T., Nellen, D., Affolter, M. & Basler, K. *Cell* **80**, 889–897 (1995).
- Letsou, A. et al. *Cell* **80**, 899–908 (1995).
- Golic, K. G. *Science* **252**, 958–961 (1991).
- Xu, T. & Rubin, G. M. *Development* **117**, 1223–1237 (1993).
- Couso, J. P., Bate, M. & Martinez-Anias, A. *Science* **259**, 484–489 (1993).
- Ingham, P. W. & Hidalgo, A. *Development* **117**, 283–291 (1993).
- Struhl, G. & Basler, K. *Cell* **72**, 527–540 (1993).
- Wilder, E. L. & Perrimon, N. *Development* **121**, 477–488 (1995).
- Diaz-Benjumea, F. J. & Cohen, S. M. *Development* **120**, 1661–1670 (1994).
- Diaz-Benjumea, F. J. & Cohen, S. M. *Development* **121**, 4215–4225 (1995).
- Niswander, L., Jeffrey, S., Martin, G. R. & Tickle, C. *Nature* **371**, 609–612 (1994).

24. Parr, B. A. & McMahon, A. P. *Nature* **374**, 350–353 (1995).
 25. Yang, Y. & Niswander, L. *Cell* **80**, 939–947 (1995).
 26. Campbell, G., Weaver, T. & Tomlinson, A. *Cell* **74**, 1113–1123 (1993).
 27. Campbell, G. & Tomlinson, A. *Development* **121**, 619–628 (1995).
 28. Blair, S. S. *Development* **115**, 21–33 (1992).
 29. Williams, J. A., Bell, J. B. & Carroll, S. B. *Genes Dev.* **5**, 2481–2495 (1991).

ACKNOWLEDGEMENTS. We thank A. Letsou, N. Perrimon and the Bloomington stock centre for fly stocks; S. Blair, S. Carroll, T. Kornberg for antibodies; the W. M. Keck Foundation neural imaging laboratory for use of their confocal microscope; J. Hudzinski for technical help; and N. Brown, E. Rulifson and S. Blair for advice. This work was supported by the NIH National Center for Research Resources and the National Cancer Institute.

CORRESPONDENCE and requests for materials should be addressed to F.M.H. (e-mail: fmhoff@maccc.wisc.edu).

A cationic channel regulated by a vertebrate intrinsic circadian oscillator

Theresa D'Souza & Stuart E. Dryer

Program in Neuroscience, Department of Biological Science, Florida State University, Tallahassee, Florida 3206-4075, USA

SECRETORY cells of the chicken pineal gland exhibit light-sensitive circadian rhythms in melatonin release that persist *in vitro*^{1–12}. Melatonin secretion is positively regulated by cyclic AMP and intracellular Ca^{2+} (refs 8, 10–15). Cyclic AMP analogues are more effective at stimulating melatonin secretion during the circadian night owing in part to increased Ca^{2+} influx at those times¹². However, this cannot be attributed to increased activity of L-type Ca^{2+} channels¹². Here we describe an unusual 40-pS cationic channel (I_{LOT}) in cultured chicken pineal cells that is permeable to Ca^{2+} and active in the night but not during the day. I_{LOT} is not voltage- or stretch-activated, it has a characteristically long open time, and its gating persists in excised inside-out patches in the absence of Ca^{2+} or cyclic nucleotides. Daily rhythms in I_{LOT} gating are also observed in previously entrained chicken pineal cells free-running under constant dark conditions. Nighttime I_{LOT} activity is not suppressed by brief light pulses.

Chicken pineal cells were cultured for five days in an incubator equipped with lights and timers and entrained to a 12:12 light–dark cycle. Recordings were made on the fifth day in culture either 4–6 h after lights on (day) or 4–6 h after lights off (night). During the night, recordings were made under infrared illumination in a dark room. Recordings during the day were made under bright visible illumination.

During the night, we often observed an unusual ionic channel that was spontaneously active in cell-attached patches when recording electrodes contained normal extracellular saline (Fig. 1a). This channel could be easily identified because of the very long open time of the unitary currents. The amplitudes of the

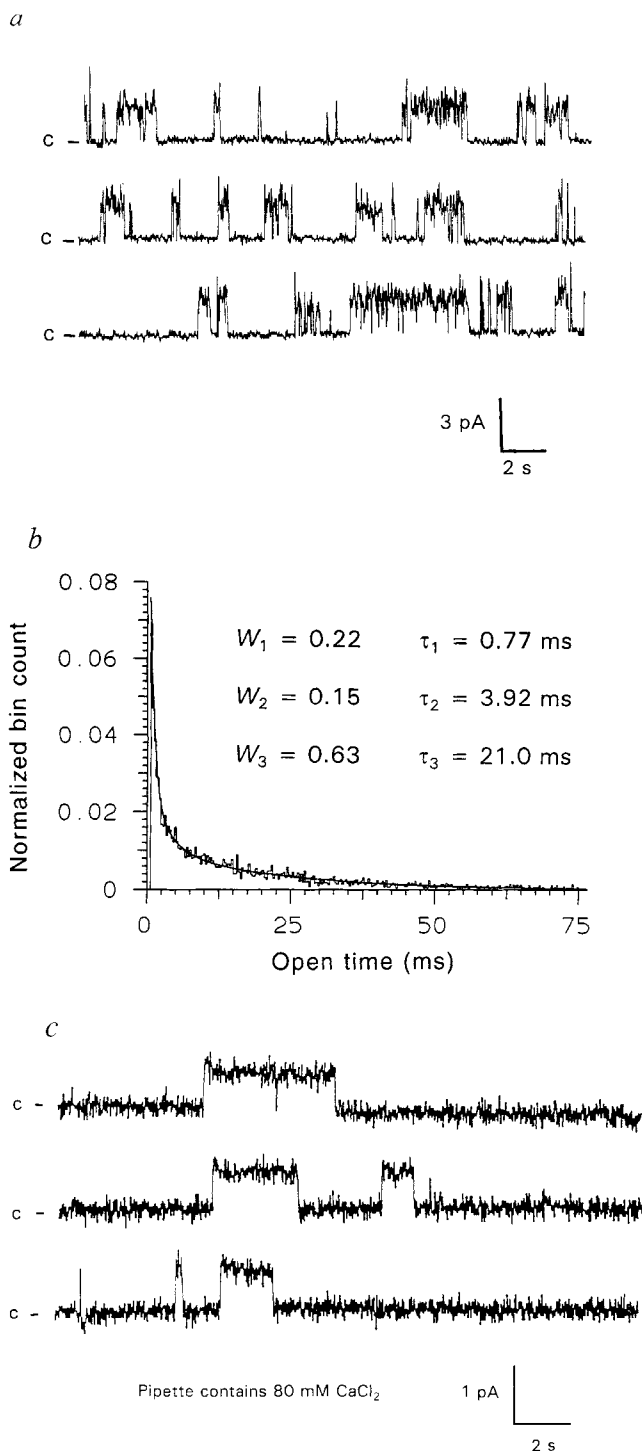


FIG. 1 Characteristics of I_{LOT} channels in cultured chick pineal cells. Recordings were made 4 h after lights-off from cells maintained for 5 d on 12:12 light–dark cycle. In all figures, currents moving out of the recording electrode are shown as outward. a, Single I_{LOT} channels in a cell-attached patch. The recording pipette was filled with normal external saline and the patch was hyperpolarized by 80 mV from the resting potential. Current level of the closed state is indicated to the left of the traces. Note that many of the bursts lasted for several seconds. b, Open-time distribution of a single I_{LOT} channel in a different cell-attached patch constructed from 10 min of continuous data. Normalized histogram is fitted with a probability distribution function containing three exponential components with weights and time constants indicated. c, I_{LOT} channels in a cell-attached patch. Conditions were as in a, except that the recording pipette was filled with a solution consisting of 80 mM CaCl_2 , 10 mM HEPES, pH 7.4.

METHODS. Chick pineal cells were dissociated as described^{16,17,19}. Cells were cultured in a medium consisting of Dulbecco's minimal essential

medium supplemented with 10% fetal calf serum, 6% chick embryo extract, 2 mM glutamine, 50 U ml⁻¹ penicillin and 50 $\mu\text{g ml}^{-1}$ streptomycin, and maintained for 5 d at 37 °C in a CO_2 incubator equipped with lights and timers set to a 12:12 light–dark cycle. On the fifth day, coverslips were transferred to a recording chamber 4 h after lights off. Recordings were made at room temperature (21–23 °C) in a dark room under infrared illumination achieved by passing a Wratten 87C filter (800-nm cutoff) into the light path of the inverted stage microscope. An infrared-sensitive videocamera (Javelin Electronics Newchip II) allowed for electrode placement under visual control. Cells were perfused continuously with an external saline consisting of 150 mM NaCl, 5.0 mM KCl, 5.4 mM CaCl_2 , 0.8 mM MgCl_2 , 5 mM D-glucose and 10 mM HEPES–NaOH, pH 7.4. Patch-clamp recordings and data acquisition and analysis were performed using PCLAMP software (Axon Instruments) as described^{16,17}. Transitions of less than 300 μs were ignored during construction of open-time histograms. All traces were filtered at 2 kHz with an 8-pole Bessel filter.

unitary currents were reduced by patch depolarization and increased by patch hyperpolarization, but the probability of channel opening did not change over a wide range of membrane potentials ($P_o = 0.31 \pm 0.06$ at a pipette potential of +100 mV, $P_o = 0.28 \pm 0.09$ at a pipette potential of 0 mV; mean $\pm$ s.e.m.; $n = 7$ cells). Gating was not altered by application of negative or positive pressure to the recording pipette.

The kinetics of these channels were complex. The open-time histogram shown in Fig. 1b could only be fitted as the sum of at least three exponential components. Similar results were obtained in three other cells. This channel occasionally stayed open without apparent interruptions for seconds at a time, although there was invariably a high level of open-channel noise. This noise may be caused by monovalent cation block of the channel occurring at a

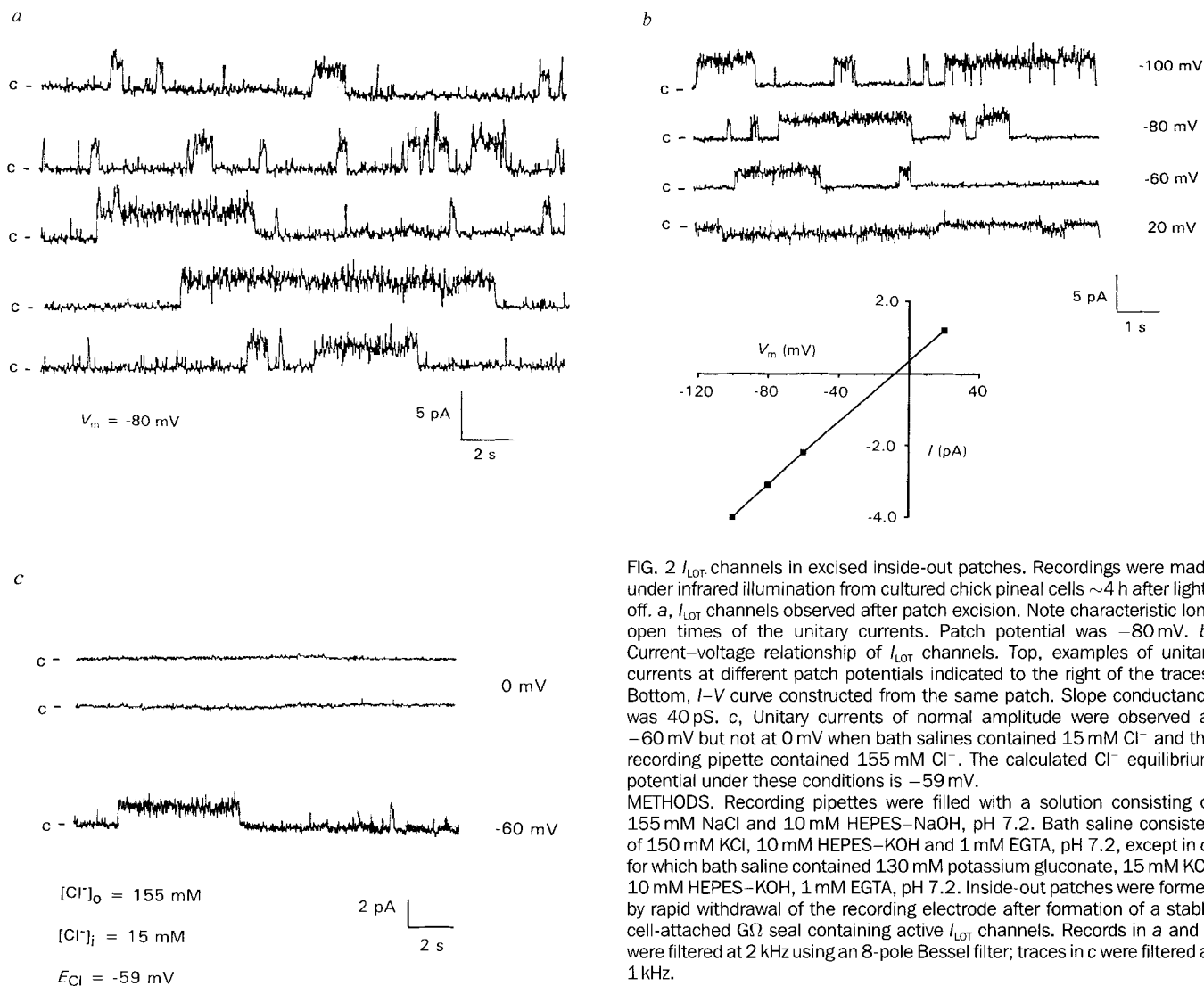
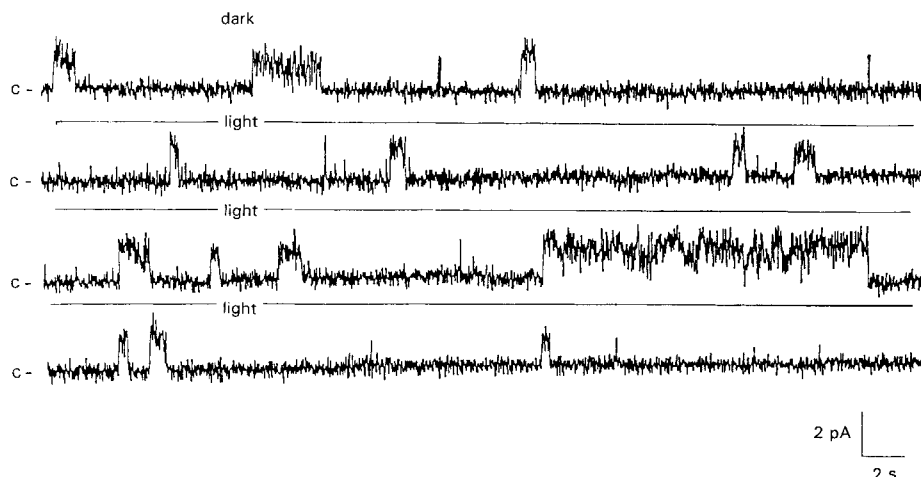


FIG. 2 I_{LOT} channels in excised inside-out patches. Recordings were made under infrared illumination from cultured chick pineal cells ~ 4 h after lights off. a, I_{LOT} channels observed after patch excision. Note characteristic long open times of the unitary currents. Patch potential was -80 mV. b, Current-voltage relationship of I_{LOT} channels. Top, examples of unitary currents at different patch potentials indicated to the right of the traces. Bottom, I - V curve constructed from the same patch. Slope conductance was 40 pS. c, Unitary currents of normal amplitude were observed at -60 mV but not at 0 mV when bath salines contained 15 mM Cl^- and the recording pipette contained 155 mM Cl^- . The calculated Cl^- equilibrium potential under these conditions is -59 mV.

METHODS. Recording pipettes were filled with a solution consisting of 155 mM NaCl and 10 mM HEPES-NaOH, pH 7.2. Bath saline consisted of 150 mM KCl, 10 mM HEPES-KOH and 1 mM EGTA, pH 7.2, except in c, for which bath saline contained 130 mM potassium gluconate, 15 mM KCl, 10 mM HEPES-KOH, 1 mM EGTA, pH 7.2. Inside-out patches were formed by rapid withdrawal of the recording electrode after formation of a stable cell-attached $G\Omega$ seal containing active I_{LOT} channels. Records in a and b were filtered at 2 kHz using an 8-pole Bessel filter; traces in c were filtered at 1 kHz.

FIG. 3 Effects of visible light on the gating of I_{LOT} channels in a cultured chick pineal cell 4 h after lights off. The cell-attached patch was hyperpolarized by 80 mV from the resting potential. Top trace shows I_{LOT} channel in cell-attached patch under infrared illumination. The next three traces show continued I_{LOT} activity in the same patch during exposure to 100-lux visible light. Traces were filtered at 2 kHz.



frequency higher than the resolution of our amplifier. The long-lived bursts account for a considerable fraction of the total integral current accumulated over a period of minutes. We therefore refer to this channel as the I_{LOT} (long open time) cationic channel. Bursts of I_{LOT} openings were often separated by quiescent periods of several seconds. I_{LOT} was also detected during the night-time in cell-attached patches when recording pipettes were filled with 80 mM CaCl_2 and 10 mM HEPES, pH 7.2 (see below). Although unitary current amplitudes and open-channel noise were reduced under those conditions, the characteristic long open times still allowed for unambiguous identification of these channels (Fig. 1c). This result indicates that I_{LOT} channels have a significant permeability to Ca^{2+} (see below).

I_{LOT} was also detected in inside-out patches. In these experiments, I_{LOT} channels were first detected in a cell-attached patch in which the pipette contained 155 mM NaCl and 10 mM HEPES, pH 7.2. The patch was then excised and the cytoplasmic face of the patch membrane was exposed to a solution containing 150 mM KCl, 10 mM HEPES and 1 mM EGTA, pH 7.2. Patch excision had no obvious effect on the gating of I_{LOT} (Fig. 2a). Mean P_o in inside-out patches was 0.25 ± 0.07 at -80 mV (analyses done in 6 patches), comparable to that observed in cell-attached patches. As with cell-attached patches, I_{LOT} gating showed little or no voltage dependence and characteristically long openings could still be observed. The unitary slope conductance of I_{LOT} under these recording conditions was 40 pS and the interpolated reversal potential of the unitary currents was -10 mV (Fig. 2b). I_{LOT} channels were not detected when five patches were held at 0 mV in bath salines containing 130 mM potassium gluconate, 15 mM KCl, 10 mM HEPES, 10 mM EGTA, pH 7.2, but were readily detected when the holding potential was switched to -60 mV, close to the calculated E_{Cl} of -59 mV (Fig. 2c). These results confirmed that I_{LOT} is a non-selective cationic channel and indicate that I_{LOT} gating is not dependent upon continued exposure to soluble cytosolic messengers^{16,17}. Moreover, application of $1 \mu\text{M}$ Ca^{2+} never caused activation of I_{LOT} in quiescent inside-out patches ($n = 21$ patches) and active I_{LOT} channels were not blocked by bath application of 1 mM MgCl_2 ($n = 6$ patches; not shown).

Active I_{LOT} channels were readily detected in cultured chick pineal cells with normal recording pipettes during the night, when melatonin secretion is highest¹⁰, but were almost never observed during the day. This was seen in cells maintained for 5 days on a 12:12 light–dark cycle, as well as in cells free-running in constant darkness for 2 days after 5 days of entrainment to 12:12 light–dark cycles (Table 1). For experiments done during 12:12 light–dark cycles, cell-attached patch recordings were made under visible illumination 4–6 h after lights-on (ZT 4–6) or under infrared illumination 4–6 h after lights-off (ZT 16–18). Recordings from free-running cells were made under infrared illumination on the second day of constant darkness 4–6 h after the expected lights-on (CT4–6) or expected lights-off (CT16–18). Statistically significant day–night differences were observed in both cases. In a separate

TABLE 1 Circadian rhythms in the gating of I_{LOT} channels

Activity of I_{LOT} channels in chick pineal cells during 12:12 LD cycle*	
Phase in LD cycle	Number of cells with active I_{LOT} channels
Light ($n = 78$ cells)	0
Dark ($n = 75$ cells)	19†
Activity of I_{LOT} channels in free-running chick pineal cells during 12:12 DD cycle‡	
Phase in circadian cycle	Number of cells with active I_{LOT} channels
Day ($n = 50$ cells)	2
Night ($n = 72$ cells)	16§

* Chick pineal cells were cultured for five days with continuous 12:12 light–dark (LD) cycles. Cell-attached patch recordings were made on the fifth day *in vitro*. Cells in the light phase were sampled 4–6 h after lights on. Cells in the dark phase were sampled 4–6 h after lights off.

† $P < 0.005$, χ^2 test for homogeneity of two independent samples.

‡ Cultured chick pineal cells were entrained to 12:12 LD cycles for five days. On the fifth day, cells were switched to 12:12 DD at the expected lights out. Cell-attached patch recordings were made in the dark on the second day of constant dark conditions. Cells in the day were sampled 4–6 h after expected lights on. Cells in the night were sampled 4–6 h after expected lights off.

§ $P < 0.025$, χ^2 test for homogeneity of two independent samples.

set of experiments using the CaCl_2 recording pipettes already described, I_{LOT} channels were observed in 7 out of 34 cells during night-time, whereas no I_{LOT} channels were observed in any of 28 cells during the daytime ($P < 0.05$, χ^2 test). For those experiments, cells were maintained on 12:12 light–dark cycles. Moreover, I_{LOT} gating detected during the night with normal recording pipettes was not inhibited by 100-lux light pulses (Fig. 3). This result, obtained in 7 cells, suggests that I_{LOT} channels are not components of a rapidly acting phototransduction cascade.

The molecular mechanisms that control the gating of this channel are unknown, but several can be excluded. For example, gating of I_{LOT} is not voltage- or Ca^{2+} -dependent and does not require continued contact with soluble cytosolic messengers. Moreover, I_{LOT} is not activated in quiescent cells by application of $10 \mu\text{M}$ melatonin ($n = 8$ cells), activation of adenylate cyclase (by application of $10 \mu\text{M}$ forskolin; $n = 14$ cells), or by depletion of internal Ca^{2+} stores (by application of $2 \mu\text{M}$ thapsigargin in Ca^{2+} -free external saline; $n = 7$ cells) (not shown). I_{LOT} gating may be controlled by covalent modification, possibly by phosphorylation/dephosphorylation of the channel molecules. Alternatively, gating may be regulated by substances that remain associated with the plasma membrane during patch excision. Finally, it is possible that I_{LOT} channel molecules are present in the plasma membrane only during the night-time.

We have identified a new cationic channel whose gating is controlled by a vertebrate intrinsic circadian oscillator. There is a previous report of a circadian clock-regulated K^+ conductance in an invertebrate system, the basal retinal neurons of the marine mollusc *Bulla gouldiana*¹⁸, but the channels underlying that conductance have not been characterized. It is possible that I_{LOT} is important in other vertebrate circadian oscillator systems, including those of retinal photoreceptors and the suprachiasmatic nucleus. □

Received 19 January; accepted 10 May 1996.

1. Deguchi, T. *Nature* **282**, 94–97 (1979).
2. Robertson, L. M. & Takahashi, J. S. *J. Neurosci.* **8**, 12–21 (1988).
3. Robertson, L. M. & Takahashi, J. S. *J. Neurosci.* **8**, 22–31 (1988).
4. Zatz, M. & Mullen, D. *Brain Res.* **450**, 137–143 (1988).
5. Zatz, M. & Mullen, D. *Brain Res.* **453**, 51–62 (1988).
6. Zatz, M. & Mullen, D. *Brain Res.* **453**, 63–71 (1988).
7. Zatz, M. et al. *Brain Res.* **438**, 199–215 (1988).
8. Zatz, M. & Mullen, D. *Brain Res.* **463**, 305–316 (1989).
9. Nikaido, S. S. & Takahashi, J. S. *Neuron* **3**, 609–619 (1989).
10. Takahashi, J. S. et al. *Rec. Prog. Hormone Res.* **45**, 279–352 (1989).
11. Zatz, M. *J. Biol. Rhyth.* **7**, 301–311 (1992).
12. Nikaido, S. S. & Takahashi, J. S. *Brain Res.* (in the press).

13. Harrison, N. & Zatz, M. *J. Neurosci.* **9**, 2462–2467 (1989).
14. Zatz, M. *Brain Res.* **477**, 14–18 (1989).
15. Zawilska, J. B. & Nowak, J. Z. *Neurosci. Lett.* **118**, 17–20 (1991).
16. Dryer, S. E. & Henderson, D. *Nature* **353**, 756–758 (1991).
17. Dryer, S. E. & Henderson, D. (1993). *J. comp. Physiol. A* **172**, 271–279 (1993).
18. Michel, S. et al. *Science* **259**, 239–241 (1993).
19. D'Souza, T. & Dryer, S. E. *Brain Res.* **656**, 85–94 (1994).

ACKNOWLEDGEMENTS. This work was supported by a grant from the Air Force Office of Scientific Research to S.E.D. and a Florida State University Neuroscience Graduate Fellowship.

CORRESPONDENCE and requests for materials should be addressed to S.E.D. (e-mail: dryer@neuro.fsu.edu).

A receptor subtype involved in neuropeptide-Y-induced food intake

Christophe Gerald*, Mary W. Walker*,
Leoluca Criscione†, Eric L. Gustafson*,
Christine Batzl-Hartmann†, Kelli E. Smith*,
Pierre Vaysse*, Margaret M. Durkin*, Thomas M. Laz*,
David L. Linemeyer*, Andrea O. Schaffhauser†,
Steven Whitebread†, Karl G. Hofbauer†,
Robert I. Taber*, Theresa A. Branchek*
& Richard L. Weinshank*

* Synaptic Pharmaceutical Corporation, 215 College Road, Paramus, New Jersey 07652-1431, USA

† Ciba-Geigy Limited, Pharmaceutical Division, Metabolic Risk Factor Department, CH-4002 Basle, Switzerland

NEUROPEPTIDE Y (NPY) is a powerful stimulant of food intake and is proposed to activate a hypothalamic 'feeding' receptor distinct from previously cloned Y-type receptors¹. This receptor was first suggested to explain a feeding response to NPY and related peptides, including NPY 2–36, that differed from their activities at the Y1 receptor². Here we report the expression cloning of a novel Y-type receptor from rat hypothalamus, which we name Y5. The complementary DNA encodes a 456-amino-acid protein with less than 35% overall identity to known Y-type receptors. The messenger RNA is found primarily in the central nervous system, including the paraventricular nucleus of the hypothalamus. The extent to which selected peptides can inhibit adenylate cyclase through the Y5 receptor and stimulate food intake in rats correspond well. Our data support the idea that the Y5 receptor is the postulated 'feeding' receptor, and may provide a new

method for the study and treatment of obesity and eating disorders.

Using an expression cloning method, in which we screened for binding of ¹²⁵I-labelled PYY (ref. 3), we isolated a rat hypothalamic cDNA encoding a new NPY/PYY/PP-receptor subtype, which we named Y5. This receptor is distinct from other Y-type receptors, known as Y1 (refs 1,4–9), Y2 (refs 1,3,10,11), Y3 (ref. 1), Y4 (or PP1)^{12–14} and PYY-preferring receptor^{15,16}. The deduced amino-acid sequence of Y5 and an alignment with the other cloned rat Y-type receptors is shown in Fig. 1. The Y5 receptor displays very low levels of transmembrane domain identity with other Y-type receptors, as well as with other known G-protein-coupled receptors (Fig. 1 legend). This low level of amino-acid identity is common within this receptor subfamily^{3,10,11,13,14}. A highly conserved human Y5 homologue has been isolated (Genbank U56079; 87% overall identity and 99% transmembrane-domain amino-acid identity with the rat Y5 receptor). A nucleotide database search identified a 99.6% match between reverse complement sequences of the human Y5 coding region and exon 1C (including its flanking sequences) located in the 5'-untranslated region of the Y1C alternate splice variant mRNA of the human Y1 receptor¹⁷. The deduced open reading frame of the reverse complement of exon 1C and its flanking sequences corresponds to the carboxy tail, transmembrane domains VII–VI and 95 amino acids into the V–VI loop of the human Y5 receptor. This suggests that the human Y1- and Y5-receptor genes could map, in opposite orientation, to the same locus on chromosome 4q. Southern blot analysis indicates that both rat and human genomes contain a single Y5 receptor gene (data not shown). The relationship between the rat Y5 and Y1 receptor genes has not been investigated.

The rat Y5 transcript was readily detected in rat brain by northern blot analysis, but not in peripheral tissues, except for a weak signal in the testis (Fig. 2a). Using *in situ* mRNA hybridization techniques to analyse rat brain sections, the rat Y5 transcript was localized to several brain regions consistent with a role in the central regulation of feeding behaviour¹⁸ (Fig. 2b, c). The Y5

TABLE 1 *In vitro* profile for cloned Y-type receptors based on negative coupling to cAMP concentration

Compound	EC ₅₀ (nM) ± s.e.m.			
	Rat Y1	Rat Y2	Rat Y4	Rat Y5
Human NPY	0.14 ± 0.02	1.2 ± 0.2	> 1,000	0.96 ± 0.19
Porcine NPY	0.15 ± 0.01	2.7 ± 0.1	> 1,000	0.66 ± 0.15
Porcine NPY 2–36	3.4 ± 0.9	1.6 ± 0.5	> 1,000	1.2 ± 0.4
Porcine NPY 3–36	110 ± 30	2.4 ± 0.5	> 1,000	2.8 ± 0.8
Porcine NPY 13–36	300 ± 70	2.2 ± 0.7	> 1,000	20 ± 3
Porcine C2-NPY	> 1,000	6.2 ± 1.3	> 1,000	290 ± 30
Porcine [Leu 31, Pro 34]NPY	0.15 ± 0.02	> 1,000	7.1 ± 2.2	1.2 ± 0.3
Human [o-Trp 32]NPY	> 1,000	> 1,000	> 1,000	45 ± 12
Human PYY	0.70 ± 0.15	0.58 ± 0.05	> 1,000	1.0 ± 0.3
Human PYY 3–36	> 1,000	0.64 ± 0.19	> 1,000	4.2 ± 1.3
Human [Pro 34]PYY	0.37 ± 0.11	> 1,000	6.0 ± 1.2	1.3 ± 0.4
Human PP	150 ± 10	> 1,000	0.037 ± 0.002	1.4 ± 0.5
Bovine PP	> 1,000	> 1,000	0.067 ± 0.016	4.6 ± 1.6
Rat PP	> 1,000	> 1,000	0.060 ± 0.007	170 ± 30
K _b (nM) ± s.e.m.				
BIBP 3226	4.2 ± 0.6	Inactive	Inactive	Inactive

Stable transfections were done to express rat Y1, Y2 or Y4 receptors in LMTK⁻ cells and rat Y5 receptors in 293 cells, using empirical evidence for receptor-dependent signal transduction. Forskolin-stimulated cAMP concentration was measured by radioimmunoassay as previously described^{3,13}. Intact cells were stimulated simultaneously with forskolin (10 μM) and peptides at concentrations ranging from 10⁻¹³ to 10⁻⁶ M. Maximum suppression of forskolin-stimulated [cAMP] produced by human NPY (a standard control in every rat Y1, Y2 and Y5 assay) was 68 ± 3% in rat Y1-transfected LMTK⁻ cells (n = 20), 91 ± 2% in rat Y2-transfected LMTK⁻ cells (n = 13) and 63 ± 4% in rat Y5-transfected 293 cells (n = 11). Maximum suppression of forskolin-stimulated [cAMP] produced by rat PP (a standard control in every rat Y4 assay) was 68 ± 3% in rat Y4-transfected LMTK⁻ cells (n = 20). Peptides with measurable EC₅₀ values relative to the standard control in each transfected system were all agonists. Concentration–response curves were generated for standard peptides with or without pretreatment with a fixed concentration of BIBP 3226 for 20 min; the corresponding ratio of EC₅₀ values was used to calculate the binding constant K_b. BIBP 3226 did not antagonize Y2, Y4 or Y5 receptor activation at concentrations as high as 1 μM. EC₅₀ and K_b values are reported as mean ± s.e.m. (n ≥ 3) and closely match K_i values measured in independent binding assays using porcine ¹²⁵I-labelled PYY for the Y1, Y2 and Y5 receptors, and rat ¹²⁵I-labelled PP for the Y4 receptor (data not shown).

mRNA was detected in abundance in the paraventricular hypothalamic nucleus (PVN) and the lateral hypothalamus (LH), areas that are densely innervated by NPY-containing neurons and have been implicated in the control of feeding behaviour. Y5 mRNA was also detected in the arcuate nucleus, a principal source of NPY-containing projections to the PVN. This localization is consistent with Y5-receptor regulation of feeding by means of postsynaptic mechanisms in the PVN and LH, as well as presynaptic mechanisms involving arcuate-paraventricular release of NPY. The distribution of Y5 mRNA in midline thalamic nuclei, which project heavily to the PVN and to the amygdala, suggests another potential role for the Y5 receptor in regulating the emotional aspect of appetitive behaviours. Using

previous reports of Y1 mRNA localization¹⁹, the main areas where Y1 and Y5 mRNA overlap are the hippocampus, the arcuate, suprachiasmatic and paraventricular nuclei of the hypothalamus, and the medial thalamic nuclei.

To compare the pharmacological profiles of the rat Y5 receptor to other known Y-type receptors, we performed *in vitro* functional assays using Y1, Y2, Y4 or Y5 stably transfected cell lines^{3,13}. The rat Y5 receptor is different from the other Y-type receptors in being activated with equal potency by human NPY, PYY and PP (Table 1). The Y5 receptor is also activated by porcine NPY, NPY2-36 and [Leu31, Pro34]NPY, but poorly stimulated by porcine C2-NPY (ref. 20) and rat PP. This profile closely resembles that reported for the Y1-like 'feeding' receptor

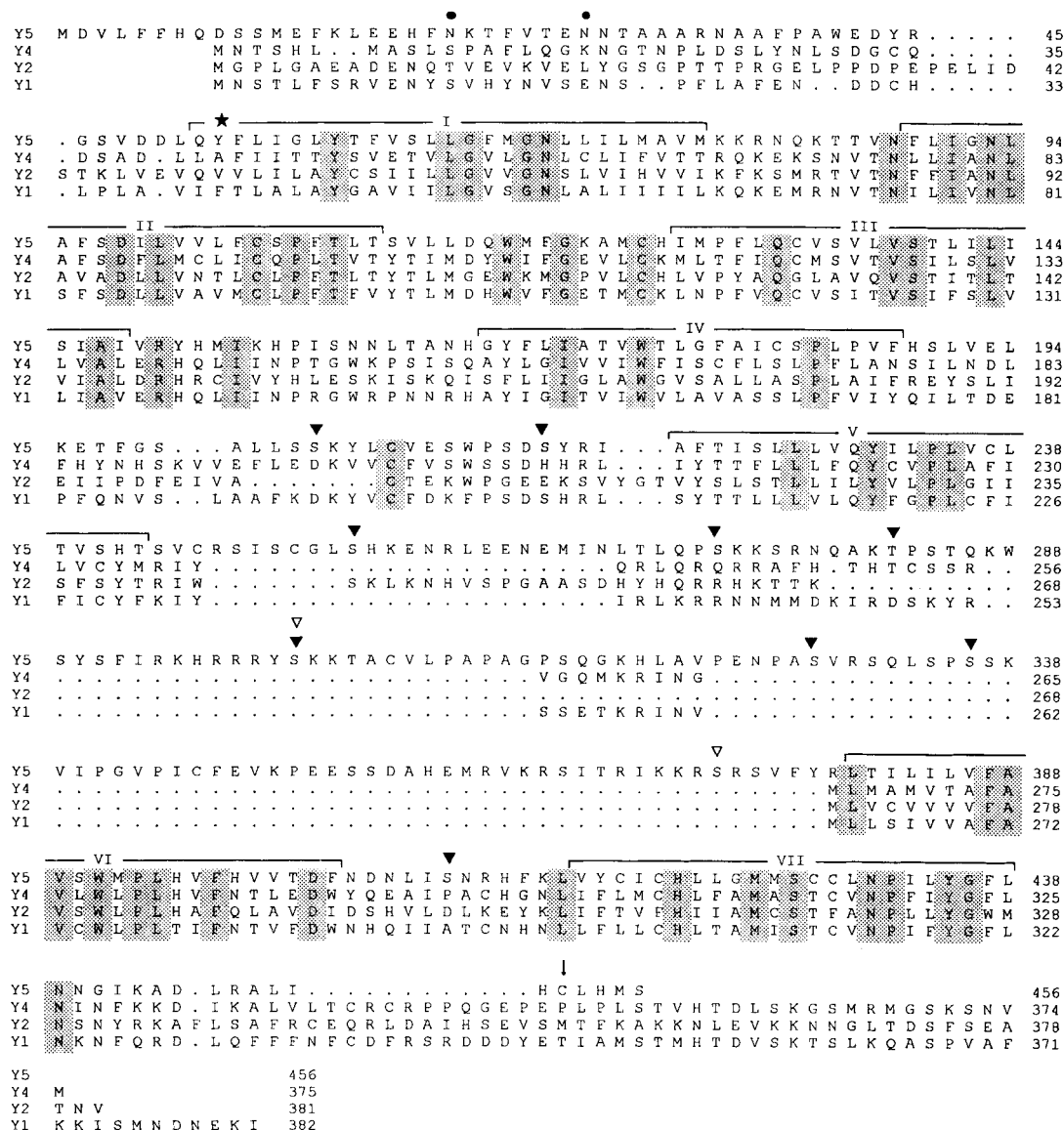
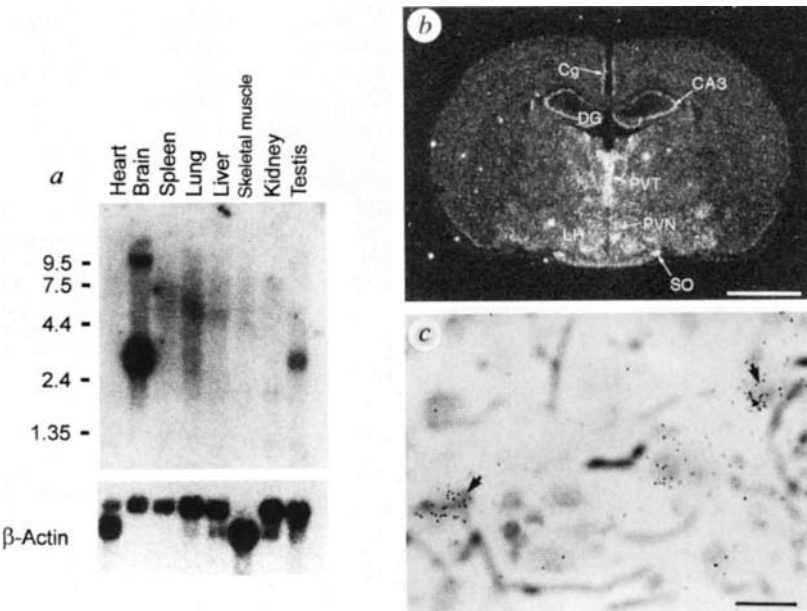


FIG. 1 Amino acid sequence corresponding to the longest open reading frame in the rat Y5 cDNA (Genbank U56078) and alignment with other known rat Y-type receptors. Amino-acid residues common to Y1, Y2, Y4 and Y5 receptors are shaded. Putative transmembrane domains I–VII, typical of G-protein-coupled receptors, are indicated above the sequence. Black circles indicate potential N-linked glycosylation sites. Black triangles correspond to the protein kinase C consensus site and white triangles to the cAMP- and cGMP-dependent protein kinase sites. The star shows a potential tyrosine kinase C phosphorylation site in transmembrane domain (TM) I and the arrow points to a possible palmitoylation site. Specific to the Y5 receptor are a 136-amino-acid V–VI loop (compared with 32, 29 and 31 amino acids for rat Y1, Y2 and Y4, respectively) and the potential

tyrosine kinase phosphorylation site in TM I. Rat Y5 amino-acid identities with rat Y1 (ref. 7), Y2 (ref. 24) and Y4 (ref. 25) are low. Overall identity 34%, 32% and 32%, respectively, and TM identities 42% in each case. The next best score is with musGIR (ref. 26) with only 32% TM identity and 27% overall identity, indicating that the rat Y5 receptor is not related to any other known G protein-coupled receptor outside of the Y-type receptor subfamily. METHODS. A rat hypothalamic cDNA library was constructed as described^{3,27} (3.4×10^6 CFU (colony-forming units); 2.7 kb insert on average). Pools of 7,500 CFU were analysed in a ¹²⁵I-labelled PYY-binding screening assay and a single clone (2.7-kb insert) corresponding to the rat Y5 receptor was isolated from a positive pool identified by direct microscopic autoradiography, as described³.

FIG. 2 a, Northern blot analysis. A strong signal is seen at 2.7 kb in whole-brain mRNA; the 8-kb band may represent unspliced hnRNA. A signal is also observed at 2.7 kb in testis. No signal was seen in other tissues after a two-day exposure. A two-hour exposure with an actin probe control is shown, demonstrating that mRNA samples were not degraded. b, Localization of Y5 mRNA by *in situ* hybridization in a coronal section through the rat brain. In the telencephalon, the most intense hybridization signals are observed over the dentate gyrus (DG), region CA3 of the hippocampus and over cingulate cortex (Cg). A 3-week exposure to Kodak XAR5 film is shown. Hybridization signals are also seen over the paraventricular thalamic nucleus (PVT) and over most hypothalamic nuclei, including the lateral hypothalamus (LH), paraventricular nucleus (PVN) and supraoptic nucleus (SO). Other positive regions include the arcuate nucleus of the hypothalamus, the central and anterior cortical amygdaloid nuclei and the dorsal raphe (data not shown). No hybridization signal was observed on sections hybridized with the radiolabelled sense oligonucleotide probe (data not shown). Scale bar, 3 mm. c, High magnification photomicrograph of the hybridization signal seen over neurons (arrows) in the paraventricular hypothalamic nucleus. A 5-week exposure to Kodak NTB2 photoemulsion is shown. Scale bar, 20 μ m.

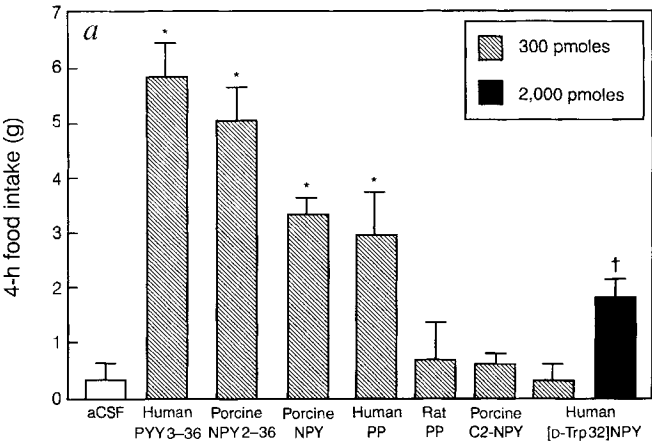
METHODS. A multiple tissue northern blot (rat MTN, Clontech, Palo Alto, CA) carrying mRNA purified from various rat tissues was hybridized at high stringency according to the manufacturer's specifications. The Y5 probe was a 0.8-kb DNA fragment generated by polymerase chain reaction, corresponding to the TM III-VI coding region of the rat Y5 receptor. The actin probe was a 2-kb human β -actin cDNA. The probes were labelled with 32 P-dATP as described³. Details of the *in situ*



hybridization method have been described previously²⁸. Sense and anti-sense 45mer oligonucleotide probes to nucleotides 1086–1130 in the V–VI loop of the rat Y5 mRNA were 3'-end-labelled with 35 S-dATP (1,200 Ci mmol⁻¹; New England Nuclear, Boston, MA), purified, and diluted in hybridization buffer.

FIG. 3 a, Peptides, selected for their ability to differentiate Y-type receptors *in vitro*, were tested at 300 pmol (ED₅₀ of NPY) for activity in feeding assays. Food intake was measured for 4 h after administration of artificial cerebral spinal fluid (aCSF), or aCSF plus peptide, and reported as mean $\pm$ s.e.m. Human PYY3–36, porcine NPY2–36, porcine NPY and human PP produced a statistically significant increase in food intake (300 pmol dose) compared with aCSF ($P < 0.01$, *). Porcine C2-NPY, rat PP and human [D-Trp 32]NPY were inactive at this dose, but the latter peptide produced a 4.5-fold stimulation of food intake when 2,000 pmol were injected ($P < 0.05$, †, $n = 30$). b, Comparison of the mean effects of peptides *in vitro* and *in vivo* indicates a role for the Y5 receptor subtype in mediating NPY-induced food intake.

METHODS. Male Sprague-Dawley rats (180–220 g) were individually housed in macrolone cages and maintained on a 11:13 h light:dark schedule (lights off at 18:00 h) at 21–23°C. Water and food (NAFAG lab chow pellets) were available *ad libitum*. Rats were implanted under pentothal anaesthesia (50 mg per kg, i.p.) with a stainless steel guide cannula targeted at the right lateral ventricle. With the incisor bar set –2.0 mm below the interaural line, stereotaxic coordinates were –0.8 mm anterior and +1.3 mm lateral to bregma. The guide cannula was placed on the dura. The injection cannula extended the guide cannula –3.8 mm ventrally to the skull surface. Animals were allowed at least 5 days of post-operative recovery. Rats were pre-screened for correct cannula placement; only those eating at least 2 g of food within 2 h after i.c.v. injection of porcine NPY (300 pmol) were retained for study. Two days later, peptides were injected in aCSF at doses of 300 or 2,000 pmol in a total volume of 5 μ l at 2 h after light onset; food intake was measured 4 h after injection. Statistical analysis was done by analysis of variance using Student-Newman-Keuls test for all peptides at the dose of 300 pmol. Student's *t*-test was used for human [D-Trp 32]NPY at the dose of 2,000 pmol on pooled data from 3 experiments.



b	In vitro potency (nM)				In vivo assay (food intake)		
	Y1	Y2	Y4	Y5	4-h food intake (g)	pmol (i.c.v.)	n
aCSF					0.4 $\pm$ 0.3	(vehicle)	6
Human PYY3–36	>1000	0.64	>1000	4.2	5.8 $\pm$ 0.6 *	300	5
Porcine NPY2–36	3.4	1.6	>1000	1.2	5.0 $\pm$ 0.6 *	300	7
Porcine NPY	0.15	2.7	>1000	0.66	3.3 $\pm$ 0.3 *	300	11
Human PP	150	>1000	0.037	1.4	2.9 $\pm$ 0.8 *	300	6
Rat PP	>1000	>1000	0.060	170	0.7 $\pm$ 0.7	300	6
Porcine C2-NPY	>1000	6.2	>1000	290	0.6 $\pm$ 0.2	300	7
Human [D-Trp 32]NPY	>1000	>1000	>1000	45	0.3 $\pm$ 0.3	300	6
Human [D-Trp 32]NPY	>1000	>1000	>1000	45	1.8 $\pm$ 0.3 †	2000	30

in rats where the rank order of potency was porcine NPY 2–36 $\geq$ porcine NPY $\approx$ porcine [Pro 34]NPY $>$ porcine C2-NPY (ref. 2). Interestingly, human [D-Trp 32]NPY (ref. 21) is a weak but selective Y5 agonist.

To test further the correlation between the Y5 receptor pharmacology and the *in vivo* feeding profile, seven peptides were selected for their ability to discriminate Y-type receptors, and administered intracerebroventricularly (i.c.v.) at a dose of 300 pmol. Food intake was monitored over four hours (Fig. 3a). Four of these peptides produced a statistically significant stimulation of food intake: human PYY 3–36, porcine NPY 2–36, porcine NPY and human PP. Rat PP, porcine C2-NPY and human [D-Trp 32]NPY were inactive at this dose. However, after injection of 2,000 pmol, the weak Y5-selective agonist [D-Trp 32]NPY did produce a small but statistically significant increase in food intake.

It is very unlikely that the rat Y1 receptor contributes to stimulation of food intake, as both human PP and human PYY 3–36 are robust stimulants of feeding but weak or inactive Y1-receptor agonists *in vitro* (Fig. 3b). In addition, food intake induced by porcine NPY (300 pmol) is not reduced by i.c.v. pretreatment with 10 nmol of the non-peptidic Y1 selective antagonist BIBP 3226 (ref. 22) (porcine NPY: 4.4 $\pm$ 0.7 g; porcine NPY + BIBP 3226: 4.0 $\pm$ 0.4 g; mean $\pm$ s.e.m.; $n = 8$). It is also unlikely that the Y2 receptor stimulates food intake, as the potent and selective Y2 agonist C2-NPY is unable to stimulate feeding. In addition, human PP is very effective at stimulating food intake, yet inactive in the Y2-receptor *in vitro* agonist assay. The fact that three orexigenic agents, porcine NPY, porcine NPY 2–36 and human PYY 3–36 are inactive at the Y4 receptor *in vitro*, and that the potent Y4 agonist, rat PP, is inactive in the feeding assay, leads us to believe that the Y4 receptor is also not involved in stimulating food intake. In contrast, the data clearly support a role for the Y5 receptor in mediating NPY-induced feeding. Strong stimulants of food intake such as human PYY 3–36, porcine NPY 2–36, porcine NPY and human PP are potent Y5 agonists *in vitro*. Conversely, peptide analogues that were inactive in the feeding assay at 300 pmol such as porcine C2-NPY, rat PP and human [D-Trp 32]NPY are also weak Y5 agonists *in vitro*. Slight discrepancies between *in vivo* and *in vitro* results might be attributed to technical aspects of i.c.v. administration of peptides, such as non-specific absorption, variable penetration to their site of action and stability in cerebrospinal fluid and brain tissue.

A restriction-site polymorphism has been described²³ in the Y1 receptor gene, 3.1 kb upstream of the Y1 coding sequence. As we report that the human Y1 and Y5 genes could map to the same locus but in opposite orientation, this polymorphism should also occur about 21 kb upstream of the Y5 coding region. This observation might have important implications and it will be interesting to determine whether this locus is associated with eating disorders or obesity in human populations.

The localization of Y5 mRNA in critical areas of hypothalamus and other brain regions also known to regulate food intake, together with an *in vitro* pharmacological profile consistent with the *in vivo* feeding data, strongly suggests that the Y5 receptor is the primary mediator of NPY-induced feeding. The isolation and characterization of the rat Y5 receptor and the identification of a human homologue will help our understanding of physiological feeding mechanisms and should facilitate the study of eating disorders. The development of selective Y5 antagonists could lead to new treatments for obesity and related disorders. $\square$

9. Blomqvist, A. G., Roubos, E. W., Larhammar, D. & Martens, G. J. M. *Biochim. biophys. Acta* **1261**, 439–441 (1995).
10. Rose, P. M. et al. *J. biol. Chem.* **270**, 22661–22664 (1995).
11. Rimland, J. M. et al. *Molec. Pharmacol.* **49**, 387–390 (1996).
12. Schwartz, T. W. *Gastroenterology* **85**, 1411–1425 (1983).
13. Bard, J. A., Walker, M. W., Branchek, T. A. & Weinschank, R. L. *J. biol. Chem.* **270**, 26762–26765 (1995).
14. Lundell, I. et al. *J. biol. Chem.* **270**, 29123–29128 (1995).
15. Laburthe, M. et al. *Endocrinology* **118**, 1910–1917 (1986).
16. Cox, H. M. & Tough, I. R. *Br. J. Pharmacol.* **116**, 2673–2678 (1995).
17. Ball, H. J., Shine, J. & Herzog, H. *J. biol. Chem.* **270**, 27272–27276 (1995).
18. Stanley, B. G. in *The Biology of Neuropeptide Y and Related Peptides* (eds Colmers, W. F. & Wahlestedt, C.) 457–510 (Humana, Totowa, New Jersey, 1993).
19. Larsen, P. J., Sheikh, S. P., Jakobsen, C. R., Schwartz, T. H. & Mikkelsen, J. D. *Eur. J. Neurosci.* **5**, 1622–1637 (1993).
20. Krstenansky, J. L., Owen, T. J., Buck, S. H., Hagaman, K. A. & McLean, L. R. *Proc. natn. Acad. Sci. U.S.A.* **86**, 4377–4381 (1989).
21. Balasubramaniam, A. et al. *J. med. Chem.* **37**, 811–815 (1994).
22. Rudolf, K. et al. *Eur. J. Pharmacol.* **271**, R11–R13 (1994).
23. Herzog, H. et al. *J. biol. Chem.* **268**, 6703–6707 (1993).
24. Gerald, C., Walker, M. W., Branchek, T. A. & Weinschank, R. L. PCT Int. Patent Application WO 95/21245 (1995).
25. Bard, J. A., Walker, M. W., Branchek, T. A. & Weinschank, R. L. PCT Int. Patent Application WO 95/17906 (1995).
26. Harrigan, M. Y., Campbell, N. F. & Bourgeois, S. *Molec. Endocr.* **5**, 1331–1338 (1991).
27. Gerald, C. et al. *EMBO J.* **14**, 2806–2815 (1995).
28. Durkin, M. M. et al. *Molec. Brain Res.* **33**, 7–21 (1995).

ACKNOWLEDGEMENTS. We thank S. Smith, C. He, S. Fried, S. Daouti, Z. Shaposhnik, N. Rogacki, T. Johnson, K. Steinberg, D. Lawrence, A. Moralishvili and Y. She for expert technical assistance, D. Eichlsberger, D. Folio and D. Mannhart for help with the animal studies, T. Schmidlin and Y. Yamaguchi for the synthesis of BIBP 3226, G. Moralishvili for graphical assistance and E. Marton for preparing the manuscript. We are grateful to E. Kandel for critical review of the manuscript.

CORRESPONDENCE and requests for materials should be addressed to C.G.

Requirement for Stat4 in interleukin-12-mediated responses of natural killer and T cells

William E. Thierfelder*, Jan M. van Deursen†, Koh Yamamoto‡, Ralph A. Tripp§, Sally R. Sarawar§, Richard T. Carson§, Mark Y. Sangster§, Dario A. A. Vignali§, Peter C. Doherty§, Gerard C. Grosveld† & James N. Ihle*¶

Departments of * Biochemistry, † Genetics, and § Immunology, St Jude Children's Research Hospital, 332 North Lauderdale, Memphis, Tennessee 38105, USA

Departments of || Pathology and ¶ Biochemistry, University of Tennessee Medical School, Memphis, Tennessee 38163, USA

‡ The First Department of Internal Medicine, Tokyo Medical and Dental University, Bunkyo-ku, Tokyo 113, Japan

SIGNAL transducers and activators of transcription (STATs) are activated by tyrosine phosphorylation in response to cytokines and mediate many of their functional responses^{1–3}. Stat4 was initially cloned as a result of its homology with Stat1 (refs 4, 5) and is widely expressed, although it is only tyrosine-phosphorylated after stimulation of T cells with interleukin (IL)-12 (refs 6, 7). IL-12 is required for the T-cell-independent induction of the cytokine interferon (IFN)- γ , a key step in the initial suppression of bacterial and parasitic infections. IL-12 is also important for the development of a Th1 response, which is critical for effective host defence against intracellular pathogens^{8,9}. To determine the function of Stat4 and its role in IL-12 signalling, we have produced mice that lack Stat4 by gene targeting. The mice were viable and fertile, with no detectable defects in haematopoiesis. However, all IL-12 functions tested were disrupted, including the induction of IFN- γ , mitogenesis, enhancement of natural killer cytolytic function and Th1 differentiation.

Received 26 April; accepted 14 June 1996.

1. Wahlestedt, C. & Reis, D. J. A. *Rev. Pharmacol. Tox.* **32**, 309–352 (1993).
2. Stanley, B. G., Magdalin, W., Seirafi, A., Nguyen, M. M. & Leibowitz, S. F. *Peptides* **13**, 581–587 (1992).
3. Gerald, C. et al. *J. biol. Chem.* **270**, 26758–26761 (1995).
4. Eva, C., Keinänen, K., Monyer, H., Seeburg, P. & Sprengel, R. *FEBS Lett.* **271**, 81–84 (1990).
5. Larhammar, D. et al. *J. biol. Chem.* **267**, 10935–10938 (1992).
6. Herzog, H. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 5794–5798 (1992).
7. Krause, J., Eva, C., Seeburg, P. H. & Sprengel, R. *Molec. Pharmacol.* **41**, 817–821 (1992).
8. Eva, C., Oberio, A., Sprengel, R. & Genazzani, E. *FEBS Lett.* **314**, 285–288 (1992).

We disrupted the *Stat4* gene by inserting a hygromycin-resistance gene cassette in the second coding exon in the opposite transcriptional orientation (Fig. 1a). Targeted ES cell clones were used to obtain chimaeric mice that transmitted the disrupted locus through the germ line. Lines from two independent clones had comparable phenotypes in our assays. Breeding of *stat4*^{+/-} mice resulted in *stat4*^{-/-} progeny (Fig. 1b) at the expected mendelian frequency among the 188 offspring examined. The *stat4*^{-/-} mice had no obvious abnormalities. *Stat4* protein was not detectable in spleen or testes extracts by immunoprecipitation and western blotting with an antiserum against the carboxy terminus of the protein, which is consistent with complete elimination of the gene product (Fig. 1c).

Stat4 is highly expressed in testes in germ cells from the mid round to the early elongating spermatid stage, when post-meiotic cells undergo the morphological changes that produce spermatozoa (ref. 4, and D. Wolgemuth, G. Herrada and J.N.I., unpublished data). However, no morphological changes in spermatogenesis were evident and homozygous males were fertile, indicating that *Stat4* fulfils a redundant function or is not essential for spermatogenesis. Similarly, there were no alterations in haematopoiesis to suggest that the developmentally regulated, myeloid expression⁴ fulfils a non-redundant, essential function. There were also no detectable changes in the percentages of major lymphoid populations expressing cell-surface markers, including CD4, CD8, CD28, CD32, CD45, CD22, B220, CD34, CD2 and CD5 (data not shown). Last, there were no detectable differences in the serum concentrations of the immunoglobulin isotypes (data not shown).

Because of the unique ability of IL-12 to induce *Stat4* tyrosine-phosphorylation, we examined IL-12 responses. One major activity of IL-12, alone or in synergy with IL-2, is induction of IFN- γ by resting and activated natural killer (NK) and T cells^{8,9}. We therefore examined the ability of splenocytes from individual, littermate, wild-type and *stat4*^{-/-} mice to produce IFN- γ (Fig. 2a).

Strikingly, stimulation of spleen cells from *stat4*^{-/-} mice with IL-12 resulted in no detectable IFN- γ production, whereas IFN- γ was readily detected in supernatants of IL-12-stimulated splenocytes from wild-type or heterozygous mice. In contrast, the production of IFN- γ in response to anti-CD3, concanavalin A, or stimulation with a combination of TPA (12-*O*-tetradecanoylphorbol-13-acetate) and ionomycin, was not detectably altered.

IL-12 was initially identified as stimulatory factor for natural killer cells that activates NK-cell-mediated cytotoxicity^{8,9}. The precise basis for this activity is not known, but may be due to IL-12-induced expression of genes for enzymes such as perforin and granzyme B¹⁰. To assess the role of *Stat4*, we examined the effects of IL-12 on NK-enriched spleen-cell populations from *stat4*^{-/-} mice. As shown in Fig. 2b, NK-enriched cell populations incubated in IL-12 had readily detectable cytolytic activity against YAC-1 NK target cells. This activity was predominantly associated with the CD5-negative population, consistent with killing by NK cells. In contrast, little NK activity was observed with cells from littermate *stat4*^{-/-} mice.

IL-12 also induces proliferation of preactivated T and NK cells^{8,9}. To investigate this activity, splenocytes from individual wild-type or *stat4*^{-/-} mice were stimulated with anti-CD3 *in vitro* for 3 days. It can be seen from Fig. 3 that splenocytes from both wild-type and *stat4*^{-/-} mice were comparably stimulated indicating the lack of any general effect of *Stat4* deficiency on lymphocyte activation by anti-CD3. Restimulation of cells from wild-type mice for 48 h with IL-12 in the presence of an antiserum against IL-2 resulted in a stimulation index of ~35-fold. However, there was no detectable response to IL-12 with activated cells from *stat4*^{-/-} mice, demonstrating a role for *Stat4* in IL-12-induced proliferation of activated lymphocytes.

Th1 and Th2 helper T cells play unique roles in immune responses^{11,12}. IL-4 facilitates differentiation of naive, CD4⁺ T cells into Th2 cells, which secrete IL-4, IL-5, IL-10 and IL-13 following antigen stimulation. This activity requires *Stat6* because

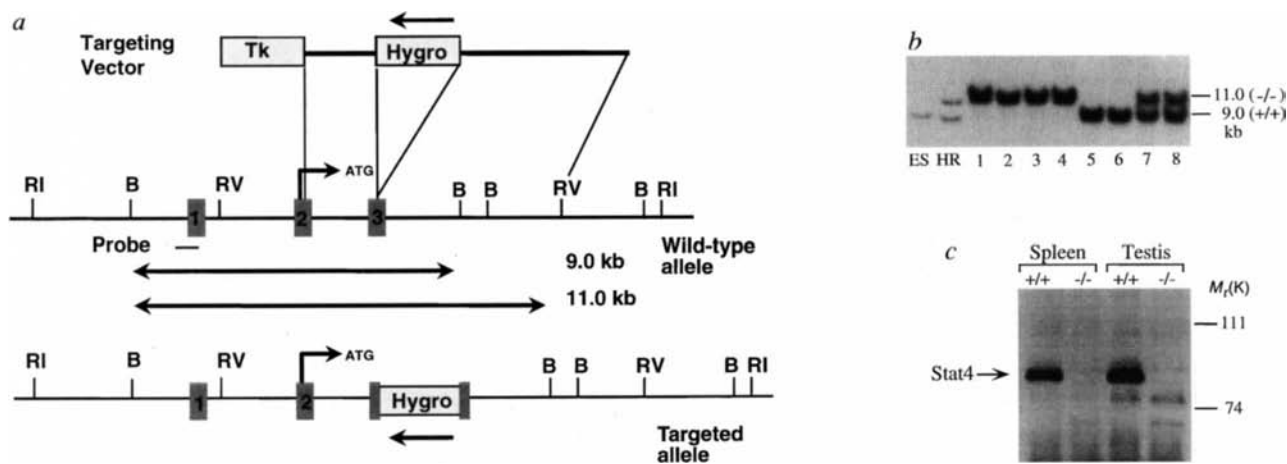


FIG. 1 Disruption of the *stat4* gene. *a*, The targeting vector (top), a *stat4* map of restriction-enzyme digestion sites (middle) and the homologous recombination (bottom) are shown. Exons 1–3 are indicated. An insertion targeting vector containing 2.0 and 4.0 kb of 5' and 3' homology was constructed for positive (Hygro) and negative selection (thymidine kinase). Sites targeted by the restriction enzymes *Bam*HI (B), *Eco*RI (RI) and *Eco*RV (RV) are indicated. *b*, Southern blot analysis of F2 DNA. *Bam*HI-digested DNAs were hybridized with a 0.45-kb *Kpn*I–*Xba*I probe. The endogenous and targeted alleles are 9.0 and 11.0 kb, respectively. Embryonic stem cell (ES) and homologous recombinant clone (HR) control DNAs are indicated, with samples from eight F2 mice (lanes 1–8). *c*, Immunoblot analysis of *Stat4* protein. Spleen cells and testes tissue were obtained from wild-type or *stat4*^{-/-} mice. *Stat4* was immunoprecipitated with a C-terminal (amino-acids 730–749) antiserum and analysed by western blotting with the same antiserum.

METHODS. *a*, Genomic sequences were cloned from a mouse strain-129

Sv/E genomic library. The hygromycin B cassette contains a thymidine kinase promoter and poly(A) site from pMC1 (Stratagene) and disrupts exon 3 by insertion into the *Mun*I site. A herpes simplex thymidine kinase cassette (HSV-tk) was used for negative selection. *b*, Electroporation into E14 ES cells and Southern blot screening for recombinant clones were done as described¹⁸. Chimaeric mice were generated by injection of four independent clones into C57BL/6 blastocysts. Heterozygotes from two clones that transmitted the targeted allele were bred separately to produce homozygous mutant (–/–) mice. *c*, Untreated spleen cells were extracted in hypotonic buffer as described¹⁹, but with 1.0 mM phenylmethylsulphonyl fluoride and without BSA. Testes were disrupted by Dounce homogenization and extracted at 50 mg ml⁻¹ in the same buffer with 1 mM dithiothreitol, 3 μ g ml⁻¹ aprotinin, 6 μ g ml⁻¹ chymostatin, 10 μ g ml⁻¹ E64, 5 μ g ml⁻¹ leupeptin and 0.7 μ g ml⁻¹ pepstatin A. Immunoprecipitation, electrophoresis and blotting were done as described¹⁹.

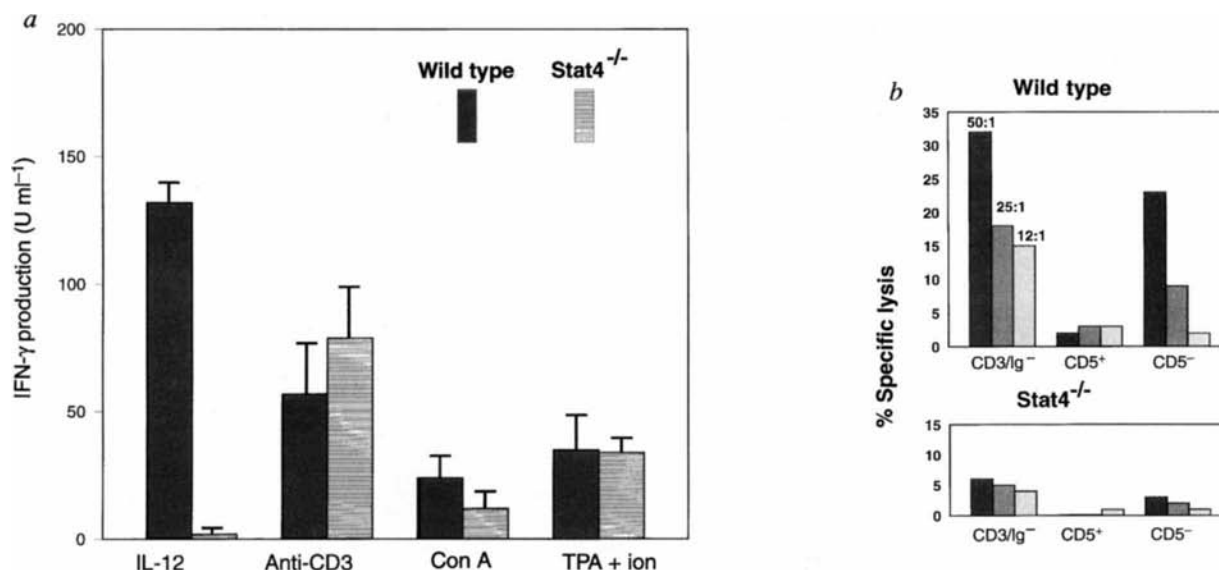


FIG. 2 Lack of IL-12-mediated IFN- γ production and NK activity in *stat4*^{-/-} mice. **a**, Lack of IFN- γ production following IL-12 stimulation of splenocytes from *stat4*^{-/-} mice. Splenocytes were obtained from either wild-type mice or *stat4*^{-/-} mice and plated in wells containing murine IL-12 p70 (20 units ml⁻¹), anti-CD3 (10 μ g ml⁻¹), Con A (2 μ g ml⁻¹), or a combination of TPA (5 ng ml⁻¹) and ionomycin (500 ng ml⁻¹) as indicated, in the presence of 50 U ml⁻¹ IL-2. Cells were cultured for 48 h and the concentration of IFN- γ in the supernatant was assayed by ELISA²⁰. The mean IFN- γ production is shown in units per ml for three wild-type and three *stat4*^{-/-} mice; the standard deviation for IFN- γ production between animals is indicated by error bars. **b**, Loss of IL-12-induced NK cytolytic activity in *stat4*^{-/-} mice. Splenic lymphocytes were obtained from wild-type or *stat4*^{-/-} mice and an NK-enriched (Ig⁻, CD3⁻) population obtained by negative selection. These cells were further fractionated into a more highly enriched NK population (CD5⁺) and residual T cells (CD5⁻) by FACS sorting. The various populations were cultured for 18 h with IL-12 (14 pM) and assayed for NK activity against ⁵¹Cr-labelled YAC-1 target cells at the indicated effector-to-target ratios.

METHODS. **a**, For stimulation with IL-12, 96-well plates were initially coated with a monoclonal antibody (C15.1, 20 μ g ml⁻¹ in PBS; gift from G. Trinchieri) against IL-12 overnight at 4 °C, washed with PBS, blocked with 5% FCS in PBS at 37 °C for 1 h and washed with PBS. IL-12 was then added at 20 U ml⁻¹ and allowed to bind for 4 h at 25 °C. For anti-CD3 stimulation plates were coated with antibody (145.2C11, 10 μ g ml⁻¹ in PBS) overnight at 4 °C and washed with PBS. Splenocytes (5 $\times$ 10⁵) were added in 0.1 ml RPMI 1640, 10% FCS to the plates. Where appropriate, TPA and ionomycin or Con A were added. For IL-12 stimulation, 50 U ml⁻¹ recombinant IL-12 (PharMingen) was also added. Plates were cultured for 48 h at 37 °C and the IFN- γ released into the medium determined by ELISA²⁰. **b**, Spleen cells were stained with biotin-conjugated anti-CD3 and rat anti-mouse Ig and depleted of T and B cells using magnetic beads (Dyna). The negatively selected population was further fractionated by FACS using a phycoerythrin-conjugated monoclonal antibody to CD5 (PharMingen). The various populations were then cultured for 18 h at 37 °C with 14 pM IL-12 in tumour cocktail medium. NK activity against ⁵¹Cr-labelled YAC-1 cells was assayed.

T cells from *stat6*^{-/-} mice fail to differentiate into Th2 cells¹³⁻¹⁵. Conversely, IL-12 enhances differentiation of Th1 cells, which produce IFN- γ . We therefore used an *in vitro* assay¹⁶ to assess T-helper-cell differentiation in *stat4*^{-/-} mice. Enriched splenic CD4⁺ T cells were cultured for 5 days to promote Th1 or Th2 differentiation, were restimulated with anti-CD3, and cytokine

production measured. Restimulation of T cells from wild-type or heterozygous mice, cultured in the presence of IL-12, resulted in production of IFN- γ (Fig. 4). However, anti-CD3 restimulation of T cells from littermate *stat4*^{-/-} mice resulted in a dramatically reduced IFN- γ production, a function that is normally not affected by the absence of Stat4 (Fig. 2a). In contrast, T cells cultured with IL-4 and antiserum against IL-12 produced little IFN- γ following stimulation with anti-CD3, but produced IL-4 as well as IL-5 and IL-10 (data not shown). This was not significantly different between littermate wild-type and *stat4*^{-/-} mice. These results therefore support a unique role for Stat4 in IL-12-promoted differentiation of Th1 cells.

The phenotype of *stat4*^{-/-} mice is interesting in several regards. The absence of an effect on spermatogenesis was surprising as Stat4 is highly expressed at a specific stage of differentiation. However, lack of any significant tyrosine phosphorylation called its role into question, as did the lack of any detectable effects on

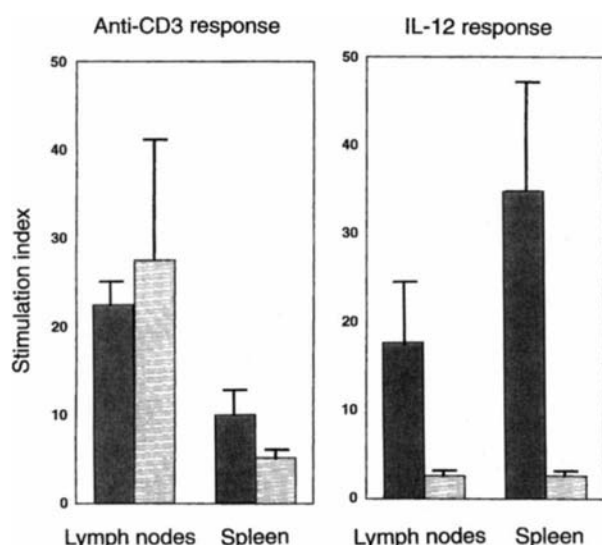


FIG. 3 Lack of a mitogenic response of activated lymphocytes from *stat4*^{-/-} mice. CD3⁺ splenic or lymph node lymphocytes were obtained by FACS sorting. Cells were cultured for three days in the presence of anti-CD3 and the extent of activation measured by tritiated thymidine incorporation for a further 24 h (left panel). In addition, aliquots of cells were centrifuged and cultured in IL-12 (100 pM) in the presence of anti-IL-2 for 24 h and the cells labelled with tritiated thymidine for a further 24 h (right panel). In each case, the extent of incorporation relative to untreated controls is plotted as the stimulation index. Results for three wild-type and three *stat4*^{-/-} mice are shown. The standard deviation for individual determinations was less than 10%.

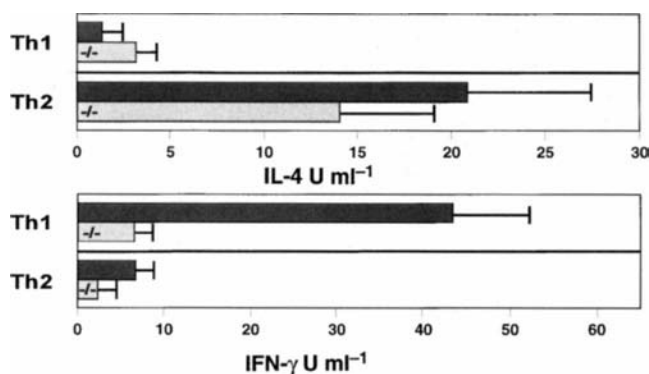


FIG. 4 *In vitro* T-helper cell differentiation. CD4⁺ T cells were stimulated under the conditions indicated for 5 days, washed and restimulated with anti-CD3 and IL-2 for 24 h. The cytokine levels in 24-h supernatants were assayed by ELISA and standardized to commercial cytokines (IL-10 from R&D; IL-4 from Genzyme, IL-5 and IFN γ from Pharmingen). Values are given as U ml⁻¹ for IL-4 or IFN γ . The detection limit for IL-4 is 0.7 U ml⁻¹ and of IL-10 is 0.1 U ml⁻¹.

METHODS. Spleen cells were cultured on plastic plates for 1 h at 37 °C to remove macrophages. Non-adherent cells were stained with monoclonal antibodies to CD8 (19.178) and class II (AFG-120.1) antigen at 4 °C for 1 h. Cells were then cultured in plates precoated with a goat anti-rat Ig and rabbit anti-mouse Ig at 4 °C for 1 h. Purified CD4⁺ cells were cultured on 20 μ g ml⁻¹ anti-CD3 ϵ (145.2C11)-coated dishes in the presence of 50 U ml⁻¹ IL-2 (Genzyme) and either anti-IL-4 (40 μ g ml⁻¹, 11B11) and IL-12 (1 ng ml⁻¹; R&D) to promote Th1 differentiation or IL-4 (50 ng ml⁻¹) to promote Th2 differentiation. All monoclonal antibodies were obtained from PharMingen. After 5 d, cells were washed and restimulated for 24 h with 1×10^6 cells in 1 ml medium containing 50 U ml⁻¹ IL-2 and anti-CD3.

spermatogenesis in *stat4*^{-/-} mice. We had also anticipated effects on haematopoiesis on the basis of expression and results¹⁷ indicating an effect of IL-12 on erythropoiesis, but again there were no detectable effects, suggesting that expression is gratuitous in both cases.

Our results demonstrate a primary, non-redundant role for Stat4 in IL-12 function, and *stat4*^{-/-} mice were deficient in all the IL-12 biological activities tested. This is surprising because IL-12 activates other signalling pathways and, for example, induces Stat3 tyrosine-phosphorylation⁶. There was also a clear effect on IL-12-induced proliferation. Whether IL-12 directly regulates cell-cycle progression or expression of cytokines or costimulatory molecules cannot be ascertained from these studies. Although little is known about the genes that are regulated by Stat4, our results suggest some candidates, including IFN- γ and enzymes associated with the cytolytic function of NK cells. The regulation of IFN- γ gene expression in activated T cells in response to a variety of stimuli was investigated, but none revealed the sequences necessary for IL-12 regulation.

Perhaps the most striking conclusions arise when considering the phenotypes of the various *stat* null mice. First, although each of the Stat proteins is widely expressed and, in the case of Stat1, is activated by several cytokines, the phenotypes of the null mice are notable for their specificity for particular cytokines. Thus, most, if not all, functions of the IFNs are lost in *stat1*^{-/-} mice, most functions of IL-4 are lost in *stat6*^{-/-} mice, and most functions of IL-12 are lost in *stat4*^{-/-} mice. The second remarkable similarity is that the functions are related to innate or acquired immunity. Indeed, it is striking that Stat4 is as critical to IL-12-enhanced Th1 differentiation as Stat6 is to IL-4-enhanced Th2 differentiation. This leads to the hypothesis that a group of cytokines, their associated receptors and the Stats, have evolved to fulfil very specific functions to enable them to survive the environmental challenges presented by invasive organisms. It will be of considerable interest to determine the phenotype of mice lacking the more widely used *stat3* and *stat5a/b* genes. □

Received 23 April; accepted 6 May 1996.

1. Ihle, J. N. *Nature* **377**, 591–594 (1995).
2. Ihle, J. N. *Cell* **84**, 331–334 (1996).
3. Schindler, C. & Darnell, J. E. Jr *Rev. Biochem.* **64**, 621–651 (1995).
4. Yamamoto, K. et al. *Molec. cell. Biol.* **14**, 4342–4349 (1994).
5. Zhong, Z., Wen, Z. & Darnell, J. E. Jr *Proc. natn. Acad. Sci. U.S.A.* **91**, 4806–4810 (1994).
6. Jacobson, N. G. et al. *J. exp. Med.* **181**, 1755–1762 (1995).
7. Bacon, C. M. et al. *J. exp. Med.* **181**, 399–404 (1995).
8. Wolf, S. F., Sieburth, D. & Sypek, J. *Stem Cells* **12**, 154–168 (1994).
9. Trinchieri, G. A. *Rev. Immun.* **13**, 251–276 (1995).
10. Salcedo, T. W., Azzoni, L., Wolf, S. F. & Perussia, B. *J. Immun.* **151**, 2511–2520 (1993).
11. Paul, W. E. & Seder, R. A. *Cell* **76**, 241–251 (1994).
12. Seder, R. A. & Paul, W. E. A. *Rev. Immun.* **12**, 635–673 (1994).
13. Shimoda, K. et al. *Nature* **380**, 630–633 (1996).
14. Takeda, K. et al. *Nature* **380**, 627–630 (1996).
15. Kaplan, M. H., Schindler, U., Smiley, S. T. & Grusby, M. J. *Immunity* **4**, 313–319 (1996).
16. Le Gros, G. et al. *J. exp. Med.* **172**, 921–929 (1990).
17. Dybedal, I., Larsen, S. & Jacobsen, S. E. W. *J. Immun.* **154**, 4950–4955 (1995).
18. van Deursen, J. et al. *Cell* **74**, 621–631 (1993).
19. Quelle, F. W. et al. *Molec. cell. Biol.* **15**, 3336–3343 (1995).
20. Sarawar, S. R. & Doherty, P. C. *J. Virol.* **68**, 3112–3115 (1994).

ACKNOWLEDGEMENTS. This research was supported by grants from the National Cancer Institute Cancer Center Support (CORE) and the National Heart, Lung and Blood Institute (to J.N.I.) and by the American Lebanese Syrian Associated Charities (ALSAC). We thank the Genetics Institute for antibody against IL-12, T. Nosaka and N. G. Jacobson for discussion, and M. Mehrpooya and C. Guglielmo for their excellent technical assistance with ELISA and FACS analysis, respectively.

CORRESPONDENCE and requests for materials should be addressed to J.N.I. (e-mail: james.ihle@stjude.org).

Impaired IL-12 responses and enhanced development of Th2 cells in Stat4-deficient mice

Mark H. Kaplan*, Ya-Lin Sun†, Timothy Hoey† & Michael J. Grusby*‡

* Department of Cancer Biology, Harvard School of Public Health, Boston, Massachusetts 02115, USA

† Tularik, Inc., 270 East Grand Avenue, South San Francisco, California 94080, USA

‡ Department of Medicine, Harvard Medical School, Boston, Massachusetts 02115, USA

INTERACTIONS between cytokine and receptor lead to the activation of multiple signalling molecules, including the family of signal transducer and activator of transcription (STAT) proteins^{1,2}. Stat4 is one member of this family^{3,4}, and is activated only in response to the cytokine interleukin(IL)-12 (refs 5, 6). By gene targeting, we have generated mice deficient in Stat4 to determine whether the function of this transcription factor is redundant with other signalling molecules activated by IL-12. IL-12-induced increases in the production of interferon(IFN)- γ cellular proliferation and natural killer (NK) cell cytotoxicity are abrogated in lymphocytes from Stat4-deficient mice. The development of Th1 cells in response to either IL-12 or *Listeria monocytogenes* is also impaired in the absence of Stat4. Furthermore, Stat4-deficient lymphocytes demonstrate a propensity towards the development of Th2 cells. These results demonstrate that Stat4 is essential for mediating responses to IL-12 in lymphocytes, and regulating the differentiation of both Th1 and Th2 cells.

D3 embryonic stem (ES) cells were transfected with a gene-targeting construct in which the exon encoding amino acids 262 to 314 of Stat4 is disrupted with a neomycin resistance gene (Fig. 1a). Mice heterozygous for the disrupted allele were intercrossed to generate Stat4-deficient (*Stat4*^{-/-}) animals and wild-type (control) littermates. All of these experiments involved mice 4–6 weeks old from the F₂ generation, with a mixture of 129 and Balb/c genetic backgrounds. Immunoblot analysis of extracts from thymus and lymph-node cells demonstrated that *Stat4*^{-/-} mice have no detect-

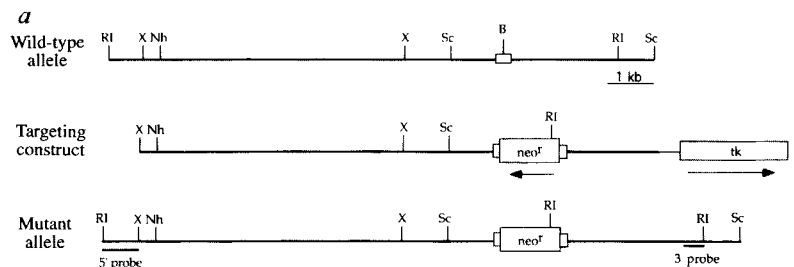
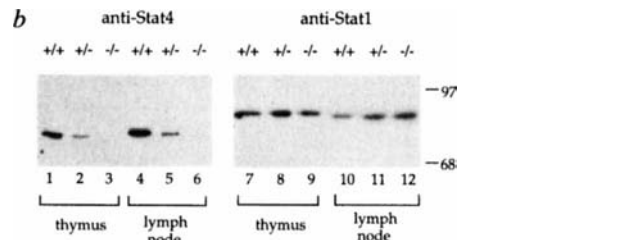


FIG. 1 Gene targeting of the *Stat4* locus. **a**, Scheme of the *Stat4* gene-targeting construct and the disrupted allele. The exon of the *Stat4* gene that was disrupted is indicated as an open box; the positions of other exons are not shown. The 5' and 3' probes used to verify correctly targeted clones are indicated under the mutant allele. Abbreviations: neo, neomycin-resistance gene; tk, herpesvirus thymidine kinase gene; B, *Bam*HI; Nh, *Nhe*I; RI, *Eco*RI; Sc, *Sal*I; X, *Xba*I. **b**, Immunoblot analysis of total cell extracts from thymus and lymph node of mice wild type, heterozygous deficient or homozygous deficient for the *Stat4* gene. As a control, filters were also probed with antisera specific for Stat1.

METHODS. D3 ES cells were transfected and mutant mice generated as described¹¹. Immunoblot analysis used antisera directed against peptides from Stat4 (Santa Cruz) or Stat1 (ref. 21), and was performed as described¹¹.



able Stat4 protein (Fig. 1b). *Stat4*^{-/-} mice are grossly indistinguishable from their wild-type littermates, and, histological examination of testes from *Stat4*^{-/-} mice failed to reveal any abnormalities despite the high level of Stat4 expression in this tissue^{3,4} (data not shown). Finally, flow cytometric analysis of lymphoid populations in the spleen, thymus and lymph nodes of *Stat4*^{-/-} mice with antibodies to CD3, CD4, CD8 and B220 showed a normal distribution of cell types (data not shown).

IL-12 stimulation of activated T lymphocytes leads to increases in IFN- γ secretion⁷. To examine the role of Stat4 in this response, spleen cells from control and *Stat4*^{-/-} mice were activated with anti-CD3 in the presence or absence of IL-12. After 24 h, supernatants were collected and IFN- γ production was quantified by enzyme-linked immunosorbent assay (ELISA). In contrast to activated cells from control mice, which show a 15-fold increase in IFN- γ secretion following IL-12 stimulation, lymphocytes from *Stat4*^{-/-} mice do not show enhanced IFN- γ secretion (Fig. 2a).

Although previous reports have suggested that Stat4 does not act directly on the IFN- γ promoter⁵, recent studies have identified a role for Stat4 in the regulation of the IFN- γ gene (X. Xu, Y.-S.L. and T.H., unpublished observations). Nevertheless, our results demonstrate that Stat4 is required for IL-12-induced IFN- γ secretion.

The involvement of STAT proteins in proliferative responses has been suggested by several studies linking these transcription factors to cellular transformation⁸⁻¹⁰, and we have recently demonstrated that Stat6 is necessary for the proliferative response of lymphocytes to IL-4 (ref. 11). To examine whether Stat4 is required for IL-12-induced proliferation, T lymphocytes were activated with anti-CD3 for four days, washed and then stimulated with IL-12 for an additional 48 h. IL-12 stimulation of activated lymphocytes from control mice leads to a dose-dependent increase in proliferation, but *Stat4*^{-/-} lymphocytes do not proliferate in response to IL-12 (Fig. 2b). IL-2 induced equivalent levels

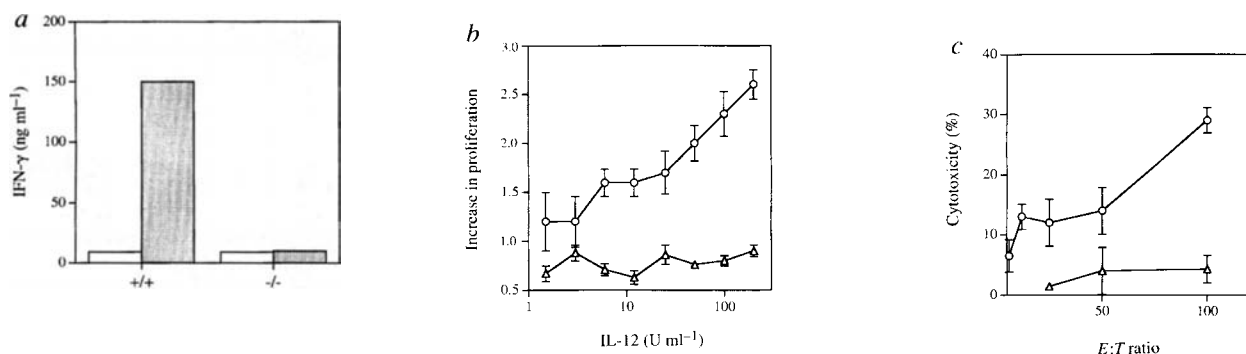


FIG. 2 IL-12-induced IFN- γ production, proliferation and NK cell cytotoxicity by control and *Stat4*^{-/-} lymphocytes. **a**, Spleen cells from control (+/+) or *Stat4*-deficient (-/-) mice were activated with anti-CD3 either alone (white bars) or with the addition of 200 U ml⁻¹ IL-12 (grey bars). Supernatants were collected after 24 h, and the amount of IFN- γ quantified by ELISA. **b**, Spleen cells from control (circles) and *Stat4*-deficient (triangles) mice were activated with anti-CD3 for 4 days and then restimulated with increasing doses of IL-12 for 48 h. Cultures were pulsed for the last 18 h with ³H-thymidine. Symbols indicate the mean of triplicate wells $\pm$ s.e.m. Proliferation in the absence of IL-12 was 43,050 $\pm$ 4,900 c.p.m. for control cells, and 90,580 $\pm$ 7,610 c.p.m. for *Stat4*^{-/-} cells. In additional assays, the proliferation of control cells in the absence of IL-12 was greater than that seen for *Stat4*^{-/-} cells, and an IL-12 dose-dependent increase in proliferation was still evident. There was no IL-12-induced proliferative response in any experiments from *Stat4*-deficient lymphocytes. **c**, Pooled

spleen cells from control (circles) or *Stat4*-deficient (triangles) mice were stimulated overnight with 200 U ml⁻¹ IL-12. Cells were collected, washed and plated with ⁵¹Cr-labelled YAC-1 cells at various E:T ratios in a 4-h chromium-release assay. Cytotoxicity from both control and *Stat4*^{-/-} cells incubated overnight without IL-12 was less than 10% at all E:T ratios. No cytotoxicity was detected against the NK-resistant target P815 by any effector-cell population with or without IL-12.

METHODS. The quantification of IFN- γ by ELISA and measurement of proliferation by ³H-thymidine incorporation were as described^{11,22}. For measurement of NK cytotoxicity, YAC-1 cells were labelled with 150 μ Ci sodium [⁵¹Cr] chromate for 2 h at 37 $^{\circ}$ C. Cells were washed three times and plated at 10⁴ cells per round-bottom well already containing appropriately diluted effector populations. After 4 h, plates were spun down and supernatants counted in a scintillation counter.

of proliferation by both control and *Stat4*^{-/-} lymphocytes (data not shown). These data establish that proliferative responses of T lymphocytes to IL-12 require Stat4.

IL-12 was first characterized by its ability to stimulate NK cell cytotoxicity¹², presumably through the induction of effector molecules such as perforin and granzymes¹³. To assess the requirement for Stat4 in this IL-12 response, we assayed for NK cell activity by using the NK-sensitive target YAC-1 and spleen-cell effectors, after incubating them overnight in the presence or absence of IL-12. In the absence of IL-12, less than 10% lysis of YAC-1 targets is seen for effector cells from both control and *Stat4*^{-/-} mice at all effector : target (E : T) ratios (data not shown). NK cells from control mice show enhanced cytotoxicity following

IL-12 stimulation, increasing to a maximum of approximately 30% specific lysis at an E : T ratio of 100 : 1 (Fig. 2c). In contrast, NK cells from *Stat4*^{-/-} mice incubated in the presence of IL-12 do not demonstrate levels of cytotoxicity above that seen from uninduced cells (Fig. 2c).

Cytokines play an important role in regulating the differentiation of T helper cells. IL-12 seems to drive the differentiation of Th1 cells^{14,15}, and IL-4 is thought to be essential for the generation of Th2 cells^{16,17}. To examine the role of Stat4 in regulating the differentiation of T helper cells, spleen cells from control and *Stat4*^{-/-} mice were differentiated into Th1 cells *in vitro* by activation with plate-bound anti-CD3 in the presence of 200 U ml⁻¹ IL-12 and 10 µg ml⁻¹ anti-IL-4, or into Th2 cells by incubation with 1000 U ml⁻¹ IL-4 and 10 µg ml⁻¹ anti-IFN-γ. After one week in culture, cells were washed and restimulated with anti-CD3, and 24 h later, supernatants were collected and analysed (by ELISA) for the presence of various cytokines. When differentiated into Th1 cells, *Stat4*^{-/-} lymphocytes produce less than 20% of the amount of IFN-γ made by control cells (Fig. 3a). A significant impairment in the production of IFN-γ could also be seen in cultures where heat-killed *Listeria monocytogenes*¹⁴ was used in place of exogenous cytokines and antibodies to drive the differentiation of Th1 cells (Fig. 3a). The production of lymphotoxin, another Th1 cytokine, is also significantly reduced in cultures of *Stat4*^{-/-} cells relative to those of control cells (data not shown). These results demonstrate that Stat4 is required to achieve substantial differentiation of Th1 cells. Indeed, the small amount of IFN-γ produced in cultures of *Stat4*^{-/-} cells may be attributed to non-Th1 cells, such as CD8⁺ or γδ⁺ T cells, or to *in vivo* preactivated cells primed by cytokines other than IL-12.

To examine the ability of *Stat4*^{-/-} lymphocytes to differentiate into Th2 cells, we quantified the levels of Th2 cytokines produced by lymphocytes cultured in the presence of IL-4 and anti-IFN-γ. Strikingly, *Stat4*^{-/-} lymphocytes produce significantly higher levels of Th2 cytokines (IL-4, IL-5 and IL-10) than control cells cultured under identical conditions (Fig. 3b). These results suggest that T cells not only fail to differentiate into Th1 cells in the absence of Stat4, but that they also demonstrate a propensity towards the development of Th2 cells. Indeed, when supernatants from lymphocytes cultured under conditions that generate Th1 cells are analysed for the presence of Th2 cytokines, significant levels of IL-5 and IL-10 can be detected in cultures of *Stat4*^{-/-} cells but not from those containing control cells (Fig. 3c). IL-4 is not detectable in cultures of Th1 differentiated cells from either control or *Stat4*^{-/-} mice (data not shown).

Our results demonstrate that Stat4 is critical for mediating IL-12-induced responses in lymphocytes. Similar requirements for Stat1 in IFN-γ-mediated responses^{18,19} and Stat6 in IL-4-mediated responses¹¹ have been shown using gene targeting. It therefore seems that, despite the activation of multiple signalling pathways by cytokines, STAT proteins are specific, non-redundant, and therefore essential, components of the cytokine signalling apparatus. Our results also demonstrate that Stat4 is important for the development of Th cell phenotypes. Stat4 not only seems to be required for promoting Th1 development, but also functions to inhibit Th2 differentiation, as seen in an increased ability of *Stat4*^{-/-} T lymphocytes to develop into Th2 cells. Under culture conditions that favour the development of Th2 cells, *Stat4*^{-/-} lymphocytes have increased Th2 cytokine secretion. Furthermore, this tendency towards Th2 development is seen even under conditions that favour the development of Th1 cells. These results suggest that, following IL-12 stimulation, Stat4 activates a genetic programme to antagonize the differentiation of Th2 cells. The demonstration that Stat4 is involved in regulating the development of both Th1 and Th2 cells is consistent with the observation that IL-12 receptor signalling becomes disabled, possibly through the loss of an IL-12 receptor component, as cells differentiate towards a Th2 phenotype²⁰. Extinction of IL-12 signalling would be a mechanism whereby a developing Th2 cell would be protected from the downregulatory effects of Stat4 activation. As the

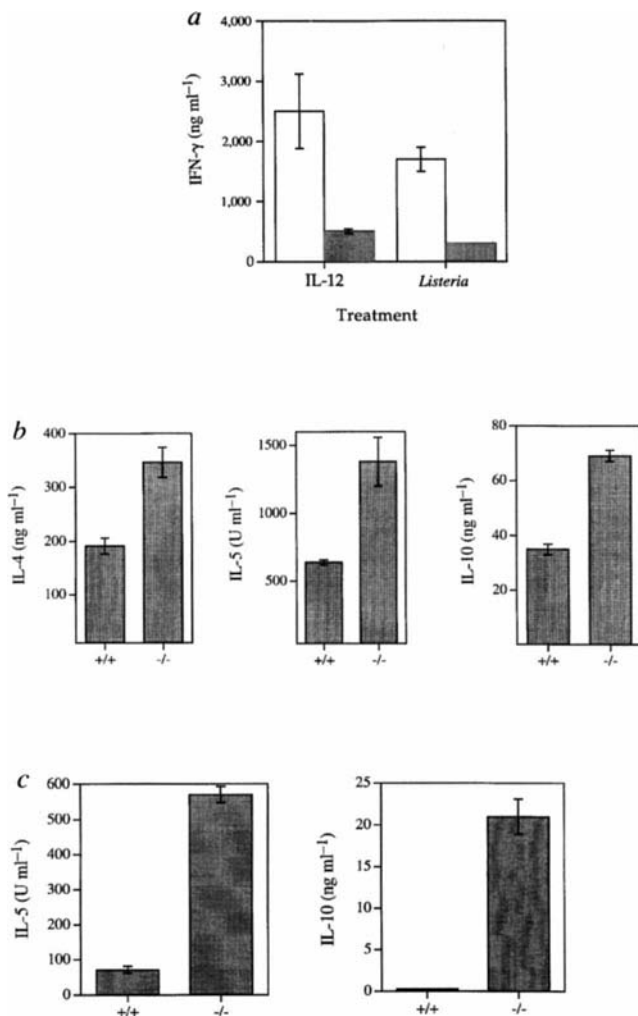


FIG. 3 Cytokine analysis of *in vitro* differentiated T lymphocytes from control and *Stat4*^{-/-} mice. **a**, Spleen cells from control (white bars) or *Stat4*^{-/-} mice (grey bars) were activated with anti-CD3 in the presence of 200 U ml⁻¹ IL-12 and 10 µg ml⁻¹ anti-IL-4 or with heat-killed *Listeria monocytogenes* for 7 days and restimulated with anti-CD3 in the absence of any exogenous antibodies and cytokines. Supernatants were collected 24 h after restimulation, and IFN-γ was quantified by ELISA as described²². **b**, Spleen cells were activated as in **a**, except that cells were differentiated into Th2 cells by incubation with 1,000 U ml⁻¹ IL-4 and 10 µg ml⁻¹ anti-IFN-γ. Supernatants were collected 24 h after restimulation, and cytokines were quantified by ELISA. **c**, Supernatants from the cultures in **a** were analysed for Th2 cytokines by ELISA. The amount of IL-10 produced by control lymphocytes was below the level of detection by ELISA (0.3 ng ml⁻¹). All of these results are representative of four independent experiments. METHODS. Cell cultures and ELISA analysis were performed as previously described^{11,12}.

differentiation of appropriate Th subsets is often essential for generating a protective immune response, our results suggest that Stat4 may be a useful target for manipulating this process. □

Received 27 March; accepted 30 May 1996.

1. Ihle, J. N. *Nature* **377**, 591–594 (1995).
2. Schindler, C. & Darnell, J. E. A. *Rev. Biochem.* **64**, 621–651 (1995).
3. Yamamoto, K. et al. *Molec. cell. Biol.* **14**, 4342–4349 (1994).
4. Zhong, Z., Wen, Z. & Darnell, J. E. *Proc. natn. Acad. Sci. U.S.A.* **91**, 4806–4810 (1994).
5. Jacobson, N. G. et al. *J. exp. Med.* **181**, 1755–1762 (1995).
6. Bacon, C. M. et al. *Proc. natn. Acad. Sci. U.S.A.* **92**, 7307–7311 (1995).
7. Chan, S. H. et al. *J. exp. Med.* **173**, 869–879 (1991).
8. Migone, T.-S. et al. *Science* **269**, 79–81 (1995).
9. Yu, C.-L. et al. *Science* **269**, 81–83 (1995).
10. Danial, N. N., Pernis, A. & Rothman, P. B. *Science* **269**, 1875–1877 (1995).
11. Kaplan, M. H., Schindler, U., Smiley, S. T. & Grusby, M. J. *Immunity* **4**, 313–319 (1996).
12. Kobayashi, M. et al. *J. exp. Med.* **170**, 827–846 (1989).
13. DeBlacker-Hohe, D. F., Yamauchi, A., Yu, C. R., Horvath-Arcidiacono, J. A. & Bloom, E. T. *Cell Immun.* **165**, 33–43 (1995).
14. Hsieh, C.-S. et al. *Science* **260**, 547–549 (1993).
15. Manetti, R. et al. *J. exp. Med.* **177**, 1199–1204 (1993).
16. Le Gros, G., Ben-Sasson, S. Z., Seder, R., Finkelman, F. D. & Paul, W. E. *J. exp. Med.* **172**, 921–929 (1990).
17. Swain, S. L., Weinberg, A. D., English, M. & Huston, G. *J. Immun.* **145**, 3796–3806 (1990).
18. Durbin, J. E., Hackenmiller, R., Simon, M. C. & Levy, D. E. *Cell* **84**, 443–450 (1996).
19. Meraz, M. A. et al. *Cell* **84**, 431–442 (1996).
20. Szabo, S. J., Jacobson, N. G., Dighe, A. S., Gubler, U. & Murphy, K. M. *Immunity* **2**, 665–675 (1995).
21. Schindler, U., Wu, P., Rothe, M., Brasseur, M. & McKnight, S. L. *Immunity* **2**, 689–697 (1995).
22. Kuchroo, V. K. et al. *Cell* **80**, 707–718 (1995).

ACKNOWLEDGEMENTS. We thank W. Paul, M. Gatley and J. Browning for antibodies, cytokines and probes, K. Sahmel for technical assistance; and L. Glimcher, S. McKnight and S. Smiley for advice and reviews of this manuscript. This work was supported by the Mathers Foundation.

CORRESPONDENCE and requests for materials should be addressed to M.J.G. (e-mail: grusby@mbcrr.harvard.edu).

Regulation of receptor-mediated endocytosis by Rho and Rac

Christophe Lamaze*, Tsung-Hsein Chuang†, Laura J. Terlecky*, Gary M. Bokoch*† & Sandra L. Schmid*

Departments of * Cell Biology and † Immunology, The Scripps Research Institute, 10666 N. Torrey Pines Road, La Jolla, California 92037, USA

PINOCTOSIS and membrane ruffling are among the earliest and most dramatic cellular responses to stimulation by growth factors or other mitogens¹. The small Ras-related G proteins Rho and Rac have a regulatory role in membrane ruffling^{1–3} and activated Rho has been shown to stimulate pinocytosis when microinjected into *Xenopus* oocytes⁴. In contrast to these well established effects of Rho and Rac on plasma membrane morphology and bulk pinocytosis, there has been no evidence for their involvement in the regulation of receptor-mediated endocytosis in clathrin-coated pits. Here we show that activated Rho and Rac inhibit transferrin-receptor-mediated endocytosis when expressed in intact cells. Furthermore, we have reconstituted these effects in a cell-free system and established that Rho and Rac can regulate clathrin-coated vesicle formation.

Receptor-mediated endocytosis of Texas red-conjugated transferrin was measured in transiently transfected cells expressing wild-type or mutant forms of Rac and Rho. Intracellular accumulation of transferrin was potently (>95% of transfected cells) inhibited in HeLa cells transiently expressing an activated, GTP-bound Rac mutant (Q61L), in which a glutamine residue is substituted for a leucine at position 61 (Fig. 1e,f), but not affected (<5%) in cells expressing either wild-type Rac (Fig. 1a,b) or an inactivated, GDP-bound mutant, Rac^{T17N} (asparagine residue

substituted for threonine at position 17; Fig. 1c,d). Similar results were obtained with activated Rho^{Q63L} (leucine instead of glutamine at position 63; Fig. 1g,h), but not with either a GDP-Rho mutant (T19N) or wild-type Rho (data not shown).

Inhibition of intracellular transferrin accumulation in transiently transfected cells could reflect Rac and Rho effects at any stage of the internalization, sorting and recycling pathway taken by transferrin receptors. Although we cannot rule out effects of Rho-family GTPases on subsequent endocytic membrane trafficking, their role in internalization was further supported by our finding that cells expressing the activating mutant of Rac (Q61L), but not wild-type or Rac^{T17N}, could not be infected with vesicular stomatitis virus (data not shown). Nevertheless, to confirm and investigate further their early role in endocytosis, we tested the effects of Rho and Rac in a cell-free system that reconstitutes the internalization of receptor-bound transferrin into clathrin-coated pits^{5,6}. Rac and Rho were both potent inhibitors (half-maximal inhibition at ~25–30 nM) of transferrin-receptor-mediated endocytosis in perforated A431 cells (Fig. 2a). The inhibitory effects of Rac and Rho must require that they be post-translationally

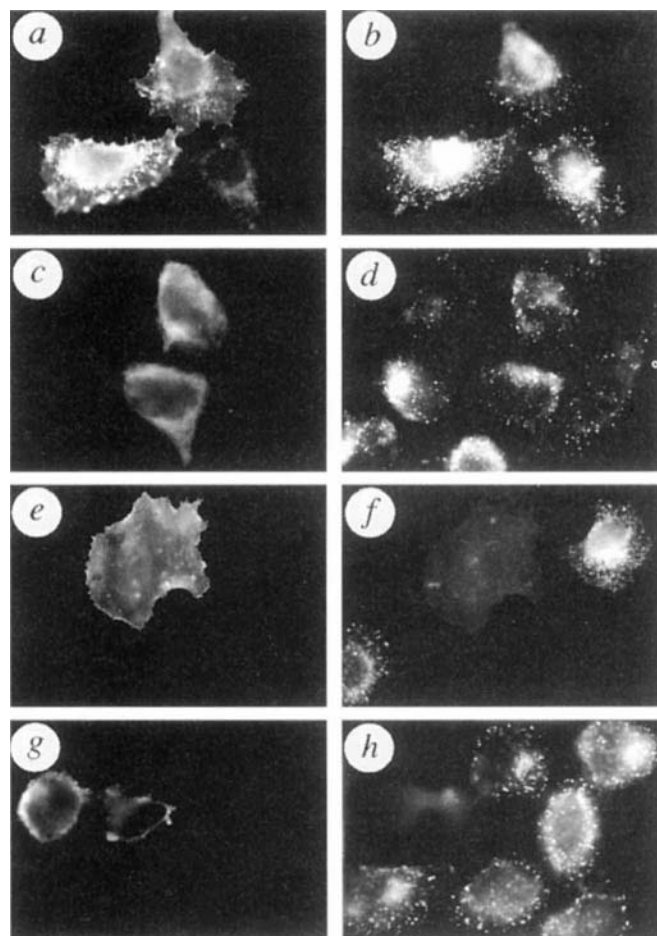


FIG. 1 Intracellular accumulation of transferrin (Tfn) is blocked in transiently transfected cells expressing activated Rac and Rho mutants. HeLa cells were grown on coverslips and transfected with expression plasmids encoding Flag-tagged wild-type Rac (a, b), Rac^{T17N} (c, d), Rac^{Q61L} (e, f) or haemagglutinin (HA)-tagged Rho^{Q63L} (g, h) as described²⁷. 24 h after transfection, cells were incubated with Texas-red-conjugated Tfn (Molecular Probes) for 15 min at 37 °C, washed and fixed with 4% paraformaldehyde (PFA) for indirect immunofluorescence as described²⁸. Internalized Tfn accumulated in punctate endosomal structures within the cytoplasm (b, d, f and h). Cells expressing Rac or Rho molecules were detected by co-immunostaining with anti-Flag (a, c, e) or anti-HA antibodies, respectively.

processed because proteins expressed in *Escherichia coli* had no effect (data not shown).

Inhibition by wild-type GTPases *in vitro* could reflect quantitative differences in the amount of protein assayed or differences in the activity of GTPase-activating protein (GAP) *in vivo* and *in vitro*. We were also unable to discern differences in the ability of Rac or Rho to inhibit endocytosis *in vitro* when preloaded either with GDP- β S or GTP- γ S, presumably because these GTPases are able rapidly to exchange nucleotides in cell lysates⁷. Therefore, to confirm the specificity of inhibition by activated forms of Rho and Rac as seen *in vivo*, assays were done in the presence of excess rhoGDI, a protein that forms a stoichiometric complex with all members of the Rho GTPase family, inhibits GDP-GTP exchange and locks them in their inactivated GDP-bound states

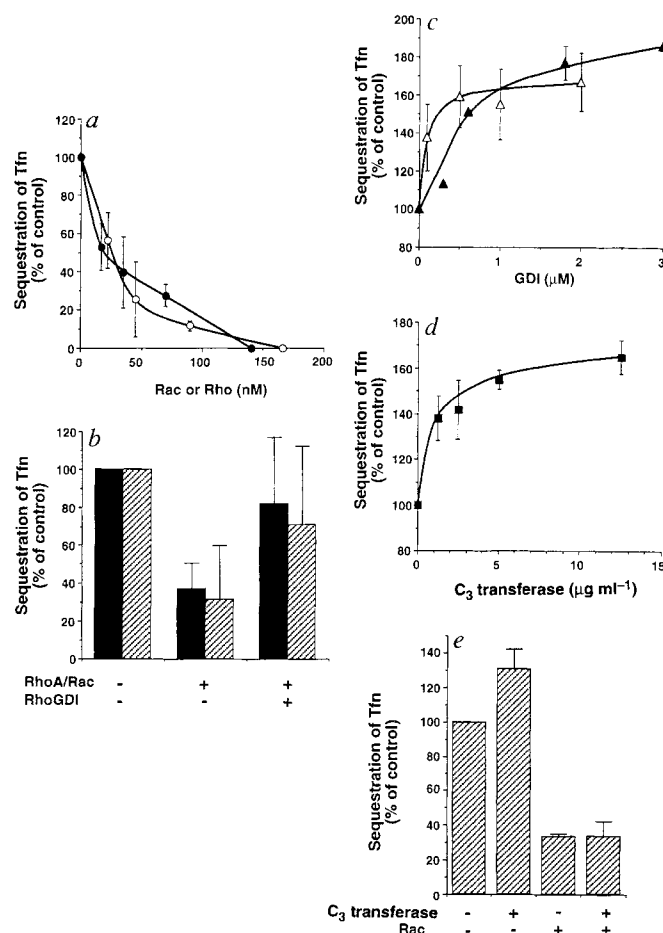


FIG. 2 Regulation by Rac and Rho of Tfn-receptor-mediated endocytosis in a cell-free system. Perforated A431 cells were incubated for 30 min at 37 °C with K562 erythroleukaemic cell cytosol, an ATP-regenerating system and biotinylated Tfn (B-Tfn). Incubations were stopped on ice and Tfn-receptor-mediated endocytosis determined by measuring the sequestration of B-Tfn from exogenously added avidin as described⁵. a, Assays were done in the presence of increasing concentrations of wild-type Rac (open circles) or Rho (filled circles) prepared from baculovirus-infected Sf9 cells as described²⁹. The efficiency of cytosol- and ATP-dependent Tfn sequestration in control incubations done in the absence of added Rho and Rac was 35 $\pm$ 8% of total cell-associated Tfn. Data shown are the average $\pm$ s.d. from 5 independent experiments. b, Assay in the presence or absence of 45 nM Rac (hatched bars) or 35 nM Rho (solid bars) with or without 1 μ M rhoGDI, prepared as described³⁰. c, Assay in the presence of increasing concentrations of purified recombinant rhoGDI (Δ) or Ly-GDI/D4 ($\blacktriangle$), or d, in the presence of a fusion protein of glutathione-S-transferase (GST) with recombinant C3 transferase (from S. Dillon, Tufts University) expressed and purified from *E. coli*. e, Assay with or without GST-C3 transferase (5 μ g ml⁻¹) and/or Rac (45 nM), as indicated. Data shown are the average $\pm$ s.d. from 3 independent experiments.

(hence GDI, GDP-dissociation inhibitor)⁸. Inclusion of rhoGDI in the assay substantially blocked the inhibitory effects of Rho and Rac (Fig. 2b), indicating that the GTP-bound forms of Rho and Rac inhibit an early event in receptor-mediated endocytosis both *in vivo* and *in vitro*.

To determine whether Rho or Rac activities are constitutively required for receptor-mediated endocytosis, perforated A431 cells were incubated in the presence of rhoGDI alone. Addition of either rhoGDI or its homologue Ly-GDI/D4 (refs 9,10) stimulated receptor-mediated endocytosis of transferrin (Fig. 2c). This result argues that Rho and Rac are not required for endocytosis and instead indicates that endogenous Rho and Rac might be negative regulators of endocytosis under the conditions of our assay. Consistent with this, recombinant C3 transferase, a bacterial toxin that specifically inactivates Rho through ADP-ribosylation of Asn41 (ref. 11) also stimulated transferrin receptor-mediated endocytosis in perforated A431 cells (Fig. 2d). Thus, activation and inactivation of Rho have opposing effects on transferrin sequestration *in vitro*. In all cases, receptor-mediated endocytosis of epidermal growth factor (EGF) was similarly affected (data not shown), indicating that Rho and Rac are general regulators of endocytosis via clathrin-coated vesicles. It should be noted that the effects seen with Rho-family GTPases were selective, because Rab-family GTPases, rabGDI or Arf1 GTPase did not have any effect on endocytosis *in vitro* (L.J.T. and S.L.S., unpublished observations).

In some cases, Rac can act through Rho in regulating downstream events^{1,12}. Thus, the Rac effects could either be mediated directly through a common downstream effector of both GTPases or, indirectly, through Rho. However, we found that C3 transferase, which specifically inactivates endogenous Rho, could not block inhibition by Rac (Fig. 2e). Thus, Rac inhibition appears to be independent of Rho.

The overall process of coated vesicle formation *in vitro* can be dissected into early events, leading to the sequestration of receptor-bound transferrin into constricted coated pits, and late events involved in coated vesicle budding⁶. Coated vesicle budding is selectively detected when transferrin, biotinylated through a cleavable disulphide bond, is internalized into sealed membrane vesicles and becomes inaccessible to the small membrane impermeant reducing agent, β -mercaptoethane sulphonic acid (MesNa). Using this assay, we found that both the rate and

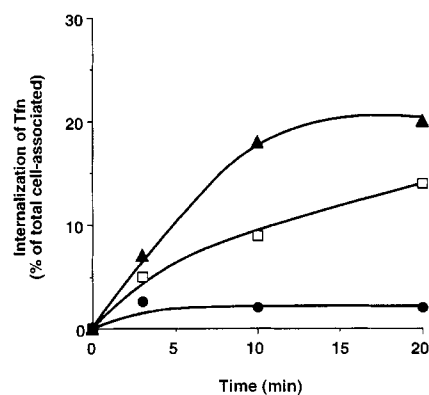


FIG. 3 Rho and rhoGDI affect clathrin-coated vesicle budding in a cell-free system. Perforated A431 cells were incubated with Tfn which had been biotinylated at a cleavable disulphide bond (BSS-Tfn) under control conditions (squares) in the presence of cytosol and ATP, or in the presence of 35 nM Rho (circles) or 1 μ M rhoGDI (triangles). Incubations were stopped on ice and the internalization of receptor-bound Tfn into sealed coated vesicles was determined by measuring the resistance of BSS-Tfn to cleavage by the small membrane-impermeant reducing agent MesNa, as described⁵. The results show ATP- and cytosol-dependent signals obtained after subtraction of background from cells incubated at 37 °C for 30 min in the presence of an ATP-depleting system and in the absence of cytosol.

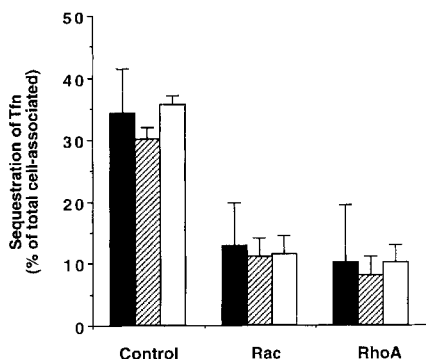


FIG. 4 Inhibition of endocytosis by Rac or Rho is not affected by either cytochalasin D or phalloidin. TfR sequestration was assayed in a cell-free system as for Fig. 2, in the absence (black bars) or presence of either 10 μM phalloidin (white bars) or 10 $\mu\text{g ml}^{-1}$ cytochalasin D (hatched bars) without (control) or with either Rac (45 nM) or Rho (35 nM), as indicated. Control experiments, both *in vivo* and *in vitro*, in which cells were fixed, permeabilized and labelled with Texas-red-phalloidin confirmed that cytochalasin D effectively disrupted actin stress fibre formation in the presence or absence of Rho and Rac (data not shown, but see ref. 12). Data are expressed as the percentage of total cell-associated ligand inaccessible to avidin after subtraction of background, as described for Fig. 3.

extent of coated vesicle budding were severely inhibited in the presence of Rho (or Rac; data not shown) and, in a reciprocal manner, were enhanced by addition of rhoGDI (Fig. 3).

Our data indicate that the effects of Rho and Rac could be mediated by a common downstream effector. As both of these small GTPases control, directly or indirectly, assembly of the actin cytoskeleton^{2,3}, we investigated whether their effects on endocytosis might be related to effects on actin by testing whether cytochalasin D, which inhibits actin assembly, or phalloidin, which stabilizes actin filaments, would antagonize or synergize with Rac and Rho. As shown in Fig. 4, neither cytochalasin D (at 1–10 $\mu\text{g ml}^{-1}$) nor phalloidin (at 0.1–10 μM) inhibit receptor-mediated endocytosis of transferrin on their own. More important, neither altered the inhibitory effects of Rac and Rho on transferrin endocytosis. Cytochalasin D treatment likewise did not affect the inhibition of transferrin accumulation seen in cells transiently transfected with either mutant Rac or Rho (data not shown). These results argue against the simple hypothesis that Rho and Rac effects are due to triggered assembly of the actin cytoskeleton. However, the effects of cytochalasin D on cortical actin assembly/disassembly are poorly characterized and we therefore cannot rule out a role for this subpopulation of actin filaments in regulating endocytosis.

The common downstream effector(s) responsible for the regulation of clathrin-mediated endocytosis by Rho and Rac have not been identified. Possible candidates include phospholipase D (PLD) (refs 13,14), phosphoinositol-4-phosphate 5-OH kinase (PI(5)K) (refs 15–17) and PI(3)K (refs 17,18) which has also been

identified as an upstream regulator of Rac and is required for growth-factor-stimulated membrane ruffling^{19–21}. However, wortmannin, a potent inhibitor of PI(3)K, has no effect either on transferrin- or EGF endocytosis in intact or perforated cells (refs 21,22; C.L. and S.L.S., unpublished observations). It remains possible that activation of PLD and/or PI(5)K might directly control clathrin-coated vesicle formation, consistent with recent proposals that site-directed lipid modifications could trigger the recruitment of components required to initiate vesicle budding^{23,24}. Alternatively, activated Rho or Rac could indirectly affect coated-vesicle formation by mistargeting, to multiple intracellular sites, the assembly of limiting components essential for coated-vesicle budding. In this context, it is interesting that GTP- γ S treatment of cells results in the mistargeting of AP2 adaptor complexes from coated pits on the plasma membrane to the endosomal compartment²⁵. Receptor-mediated endocytosis *in vitro*, only the second cell-free assay system in which the effects of Rho and Rac have been reconstituted, should provide a functional means for identifying the relevant downstream effectors.

In summary, our results are strong evidence for a previously unsuspected role for Rho and Rac in the regulation of receptor-mediated endocytosis through coated pits. This is in marked contrast to the stimulatory effects of Rac and Rho on fluid-phase endocytosis in mammalian cells¹ and *Xenopus* oocytes⁴, respectively. Our findings provide further support for the mechanistic distinction between clathrin-mediated endocytosis and clathrin-independent pinocytic events²⁶. □

Received 28 February; accepted 28 May 1996.

1. Ridley, A. J. & Hall, A. *Cell* **70**, 401–410 (1992).
2. Hall, A. A. *Rev. Cell Biol.* **10**, 31–54 (1994).
3. Takai, Y., Sasaki, T., Tanabe, K. & Nakanishi, H. *Trends biochem. Sci.* **20**, 227–231 (1995).
4. Schmalzing, G. et al. *J. Cell Biol.* **130**, 1319–1332 (1995).
5. Smythe, E., Redelmeier, T. E. & Schmid, S. L. *Meth. Enzym.* **219**, 223–234 (1992).
6. Schmid, S. L. *Trends Cell Biol.* **3**, 145–148 (1993).
7. Bokoch, G. M., Boh, B. P. & Chuang, T.-H. *J. Biol. Chem.* **269**, 31674–31679 (1994).
8. Ueda, T., Kikuchi, A., Ohgo, N., Yamamoto, J. & Takai, Y. *J. Biol. Chem.* **265**, 9373–9380 (1990).
9. Lelias, J. M. et al. *Proc. natn. Acad. Sci. U.S.A.* **90**, 1479–1483 (1993).
10. Scherle, P., Behrens, T. & Staudt, L. M. *Proc. natn. Acad. Sci. U.S.A.* **90**, 7568–7572 (1989).
11. Aktories, K. & Hall, A. *Trends pharmac. Sci.* **10**, 415–418 (1989).
12. Nobes, C. D. & Hall, A. *Cell* **81**, 53–62 (1995).
13. Malcolm, K. C., Ross, A. H., Qui, R.-G., Symons, M. & Exton, J. H. *J. Biol. Chem.* **269**, 25951–25954 (1994).
14. Singer, W. D., Brown, H. A., Bokoch, G. M. & Sternweis, P. C. *J. Biol. Chem.* **270**, 14944–14950 (1995).
15. Chong, L. D., Traynor-Kaplan, A., Bokoch, G. M. & Schwartz, M. A. *Cell* **79**, 507–513 (1994).
16. Hartwig, J. H. et al. *Cell* **82**, 643–653 (1995).
17. Tolias, K. F., Cantley, L. C. & Carpenter, C. L. *J. Biol. Chem.* **270**, 17656–17659 (1995).
18. Bokoch, G. M., Vlattos, C. J., Wang, Y., Knaus, V. G. & Traynor-Kaplan, A. E. *Biochem. J.* **315**,

- 775–779 (1996).
19. Kotani, K. et al. *EMBO J.* **13**, 2313–2321 (1994).
20. Wennström, S. et al. *Curr. Biol.* **4**, 385–393 (1994).
21. Claque, M. J., Thorpe, C. & Jones, A. T. *FEBS Lett.* **367**, 272–274 (1995).
22. Lamaze, C. & Schmid, S. L. *J. Cell Biol.* **129**, 47–54 (1995).
23. Liscovitch, M. & Cantley, L. C. *Cell* **81**, 659–662 (1995).
24. DeCamilli, P., Emr, S. D., McPherson, P. S. & Novick, P. *Science* **271**, 1533–1539 (1996).
25. Seaman, M. N. J., Ball, C. L. & Robinson, M. S. *J. Cell Biol.* **123**, 1093–1105 (1993).
26. Lamaze, C. & Schmid, S. L. *Curr. Opin. Cell Biol.* **7**, 573–580 (1995).
27. Zhang, S. et al. *J. Biol. Chem.* **270**, 23934–23936 (1995).
28. Damke, H., Baba, T., van der Bliek, A. M. & Schmid, S. L. *J. Cell Biol.* **131**, 69–80 (1995).
29. Xu, X., Barry, D. C., Settleman, J., Schwartz, M. A. & Bokoch, G. M. *J. Biol. Chem.* **269**, 23569–23574 (1994).
30. Chuang, T.-H., Xu, X., Knaus, U. G., Hart, M. J. & Bokoch, G. M. *J. Biol. Chem.* **268**, 775–778 (1993).

ACKNOWLEDGEMENTS. We thank L. Cantley for discussion. This work was supported by NIH grants to S.L.S. and G.M.B. C.L. is supported by the USAMRMC. T.H.C. is supported by a fellowship from the National Arthritis Foundation. S.L.S. is an Established Investigator of the American Heart Association.

CORRESPONDENCE and requests for materials should be addressed to S.L.S. (e-mail: slschmid@scripps.edu).

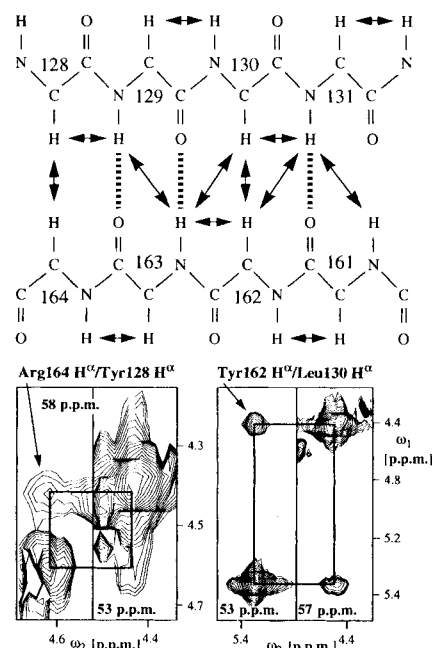
NMR structure of the mouse prion protein domain PrP(121–231)

Roland Riek, Simone Hornemann, Gerhard Wider, Martin Billeter, Rudi Glockshuber & Kurt Wüthrich

Institut für Molekularbiologie und Biophysik, Eidgenössische Technische Hochschule-Hönggerberg, CH-8093 Zürich, Switzerland

The 'protein only' hypothesis¹ states that a modified form of normal prion protein triggers infectious neurodegenerative diseases, such as bovine spongiform encephalopathy (BSE), or Creutzfeldt–Jakob disease (CJD) in humans^{2,4}. Prion proteins are thought to exist in two different conformations⁵: the 'benign' PrP^C form, and the infectious 'scrapie form', PrP^{Sc}. Knowledge of the three-dimensional structure of PrP^C is essential for understanding the transition to PrP^{Sc}. The nuclear magnetic resonance (NMR) structure of the autonomously folding PrP domain comprising residues 121–231 (ref. 6) contains a two-stranded antiparallel β -sheet and three α -helices. This domain contains most of the point-mutation sites that have been linked, in human PrP, to the occurrence of familial prion diseases⁷. The NMR structure shows that these mutations occur within, or directly adjacent to, regular secondary structures. The presence of a β -sheet in PrP(121–231) is in contrast with model predictions of an all-helical structure of PrP^C (ref. 8), and may be important for the initiation of the transition from PrP^C to PrP^{Sc}.

The NMR structure of PrP(121–231) (Fig. 1a and Table 1) contains three α -helices and a two-stranded antiparallel β -sheet. The approximate lengths of the helices are from residues 144 to 154, 179 to 193, and 200 to 217, and the lengths of the β -strands are from residues 128 to 131, and 161 to 164. The first turn of the second helix and the last turn of the third helix are linked by the single disulphide bond in the protein. The twisted V-shaped arrangement of these two longest helices forms the scaffold onto which the short β -sheet and the first helix are anchored. At the present stage of refinement, all regular secondary-structure elements and the connecting loops are well defined (see Table 1 and Fig. 1d, e), with the sole exception of residues 167 to 176. The polypeptide fold is stabilized by hydrophobic interactions in a core that contains side chains of the second helix (residues 179, 180 and 184), the third helix (residues 203, 206, 209, 210, 213 and 214), the β -sheet (Val 161), the first, mostly hydrophilic, helix (Tyr 150), and three loop regions (residues 134, 137, 139, 141, 157, 158 and 198). With the exceptions of Ile 139, Ile 184 and Val 203, the residues of the hydrophobic core are invariant in the known mammalian prion protein sequences⁹ (Fig. 3a). Hydrophobic surface patches in PrP(121–231) are located near the β -sheet and the loop preceding the first helix. The surface of PrP(121–231) is otherwise characterized by a markedly uneven distribution of positively and negatively charged residues (Fig. 1b, c).



Mature mouse PrP^C is a glycosylated 208-residue protein (codons 23–231, with deletion of codon 55 (ref. 9)) that is attached to the cell surface by means of a glycosyl phosphatidyl inositol anchor at its carboxy-terminal Ser 231 (ref. 10). It seems to be necessary for normal synaptic function¹¹ (but see ref. 12), long-term survival of Purkinje neurons¹³, and the regulation of circadian activity rhythms and sleep¹⁴. The segment of residues 109–218 was predicted to form a four-helix-bundle with helices at residues 109–122, 129–141, 178–191 and 202–218 (ref. 8), whereas prediction algorithms did not yield conclusive results for the segment 23–108, which contains the prion-characteristic octapeptide repeats. Attempts to express PrP(108–231) in the periplasm of *Escherichia coli* resulted in proteolytic cleavage after residues 112, 118 and 120, whereas PrP(121–231) is stable against degradation in *E. coli*, folds cooperatively and reversibly at pH 7 ($\Delta G_{\text{Fold}} = -22 \text{ kJ mol}^{-1}$), and is soluble at 1 mM concentration in distilled water between pH 4.0 and pH 8.5 (ref. 6). This segment also contains six of nine point-mutation sites in mature PrP that have been associated with familial prion diseases⁷ (Fig. 3a), as well as both glycosylation sites of PrP and its single disulphide bond¹⁰. We therefore chose to use PrP(121–231) for the present NMR structure determination. This choice was also supported by the demonstration that the segment 81–231 of mouse PrP is sufficient for propagation of the prion disease *in vivo*¹⁵, indicating that the C-terminal part of PrP is of special functional importance.

In the context of the structure predictions for PrP^C (ref. 8), the β -sheet in the NMR structure of PrP(121–231) is an unexpected feature. Evidence for the identification of the β -sheet is shown in

TABLE 1 Parameters characterising the NMR structure determination of PrP(121–231)

Extent of assignments (backbone and side chain ^1H , ^{13}C , backbone ^{15}N)	93%
Number of distance constraints	1,368
Number of dihedral angle constraints	227
Distance constraint violations $>0.1 \text{ \AA}$ (per conformer)	1.5 ± 1.3
Dihedral angle constraint violations $> 2.5^\circ$ (per conformer)	0.15 ± 0.36
Intra-protein AMBER energy (kcal mol^{-1})	$-5,041 \pm 97$
R.m.s.d. to the mean for N, C $^\alpha$ and C $^\beta$ of residues 125–166 and 177–219	$1.4 \text{ \AA}$
R.m.s.d. to the mean for all heavy atoms of residues 125–166 and 177–219	$2.0 \text{ \AA}$

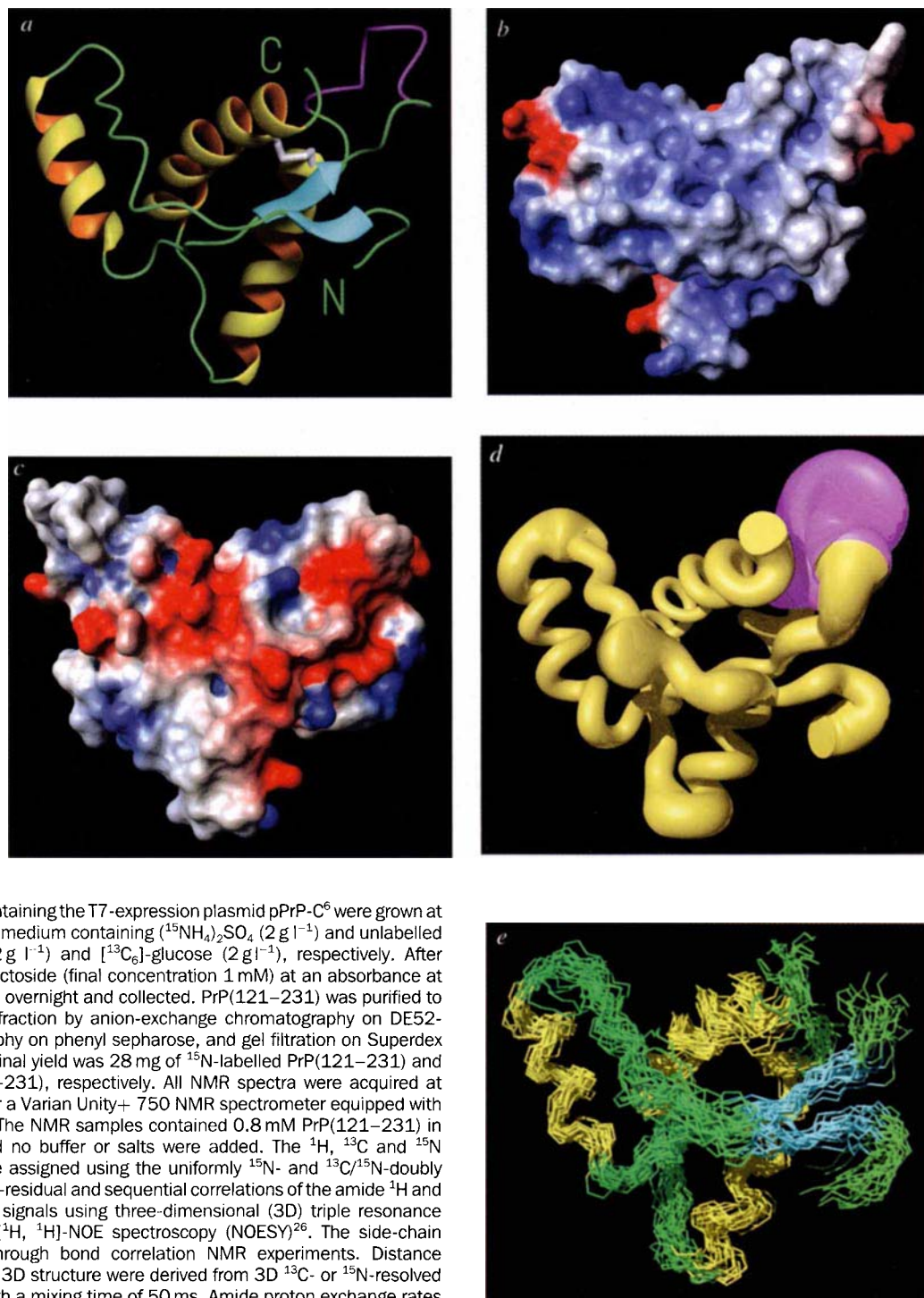
The NMR structure of PrP(121–231) was calculated with the program DIANA²¹. Starting from 100 randomized structures, the 20 conformers with the lowest DIANA target function values were energy minimized in a water shell of 6 $\text{\AA}$ minimal thickness, using the program OPAL (P. Luginbühl, P. Güntert, M. Billeter and K. Wüthrich, submitted) with the AMBER force field²².

◀ FIG. 2 Identification of the antiparallel β -sheet in PrP(121–231). Top, β -sheet with arrows showing the NOEs from which this structure was identified before the structure calculation¹⁶. Broken lines indicate β -sheet hydrogen bonds for which slowed amide–proton exchange was observed. Bottom, spectral regions from a 3D ^{13}C -correlated ^1H , ^1H -NOESY spectrum recorded in D_2O solution; squares connect the cross peaks and diagonal positions for the $d_{xx}(128, 164)$ and $d_{xx}(130, 162)$ connectivities.

Fig. 2, including the raw data on the two d_{xx} interstrand nuclear Overhauser enhancement (NOE) connectivities, which are the most direct NMR identifiers of antiparallel β -structure¹⁶. Considering the proposed increase of the β -sheet content in PrP from PrP^C to PrP^{Sc} (refs 5,17), it is tempting to speculate that the short β -sheet shown in Fig. 2 might be a 'nucleation site' for a conformational transition that could include the loops connecting the β -sheet to the first helix, which is predominantly hydrophilic and does not show amphipathic character. A systematic search of the Brookhaven data bank¹⁸ with the program DALI¹⁹ did not lead to the identification of other

FIG. 1 Globular fold and surface properties of PrP(121–231). *a*, Ribbon diagram of the structure of the mouse prion protein domain PrP(121–231), indicating the positions of the three helices (yellow) and the antiparallel two-stranded β -sheet (cyan). The connecting loops are displayed in green if their structure is well defined, and in magenta otherwise. The disulphide bond between Cys 179 and Cys 214 is shown in white. The N-terminal segment of residues 121–124 and the C-terminal segment 220–231 are disordered and not displayed. *b*, *c*, Surface of the structure of PrP(121–231). Colours indicate the electrostatic potential²³, with blue for positive charges, red for negative charges. *b*, Same orientation as *a*. *c*, View after 180° rotation about a vertical axis. *d*, *e*, Representations of the precision of the structure determination. *d*, Display of the backbone of PrP(121–231) as a cylindrical rod of variable radius, which represents the global displacements among the 20 conformers used to represent the NMR structure (Table 1)²⁴. Same orientation as *a*. Yellow represents residues 125–166 and 177–219, and the disordered loop of residues 167–176 is shown in magenta. *e*, Superposition of the 20 conformers; colours as in *a*.

METHODS. For the production of uniformly ^{15}N -labelled and $^{15}\text{N}/^{13}\text{C}$ -doubly labelled PrP(121–231), cells of *E. coli* BL21(DE3)²⁵ containing the T7-expression plasmid pPrP-C⁶ were grown at room temperature in 10 l of minimal medium containing $(^{15}\text{NH}_4)_2\text{SO}_4$ (2 g l^{-1}) and unlabelled glucose (5 g l^{-1}) or $(^{15}\text{NH}_4)_2\text{SO}_4$ (2 g l^{-1}) and $[^{13}\text{C}_6]\text{-glucose}$ (2 g l^{-1}), respectively. After induction with isopropyl- β -D-thiogalactoside (final concentration 1 mM) at an absorbance at 550 nm of 0.7, the cells were grown overnight and collected. PrP(121–231) was purified to homogeneity from the periplasmic fraction by anion-exchange chromatography on DE52-cellulose, hydrophobic chromatography on phenyl sepharose, and gel filtration on Superdex 200, as described elsewhere⁶. The final yield was 28 mg of ^{15}N -labelled PrP(121–231) and 16 mg of $^{13}\text{C}/^{15}\text{N}$ -labelled PrP(121–231), respectively. All NMR spectra were acquired at 20 °C either on a Bruker AMX 600 or a Varian Unity+ 750 NMR spectrometer equipped with triple-resonance z-gradient probes. The NMR samples contained 0.8 mM PrP(121–231) in 90% $\text{H}_2\text{O}/10\%$ D_2O at pH 4.5, and no buffer or salts were added. The ^1H , ^{13}C and ^{15}N resonances with the backbone were assigned using the uniformly ^{15}N - and $^{13}\text{C}/^{15}\text{N}$ -doubly labelled proteins by establishing intra-residual and sequential correlations of the amide ^1H and ^{15}N resonances with C^α , C^β and H^γ signals using three-dimensional (3D) triple resonance experiments and 3D ^{15}N -resolved ^1H , ^1H -NOE spectroscopy (NOESY)²⁶. The side-chain signals were assigned from 3D through bond correlation NMR experiments. Distance constraints for the calculation of the 3D structure were derived from 3D ^{13}C - or ^{15}N -resolved ^1H , ^1H -NOESY spectra recorded with a mixing time of 50 ms. Amide proton exchange rates were measured by recording a series of $[^{15}\text{N}$, $^1\text{H}]$ -correlation spectroscopy (COSY) experiments immediately after dissolving lyophilized PrP(121–231) in D_2O . The program MOLMOL²⁴ was used to generate the figure.



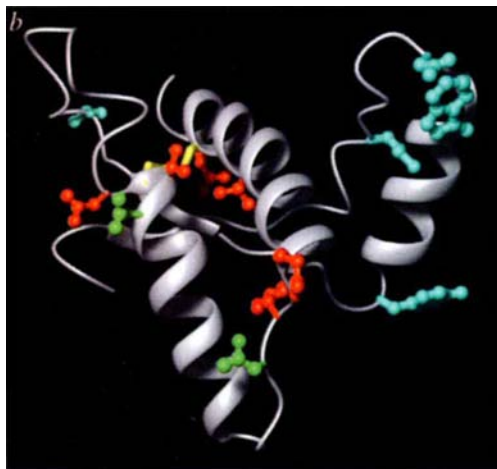
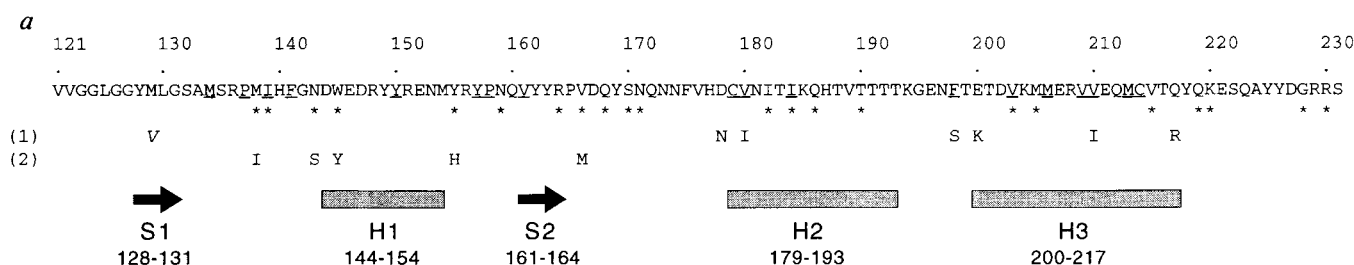


FIG. 3 Location in the 3D structure of PrP(121–231) of residues involved in sequence variations among mammalian prion proteins, and of residues that have been associated with the species barrier of prion disease transmission and with inherited prion diseases. **a**, Sequence and regular secondary structure of mouse PrP(121–231). Residues contributing to the hydrophobic core of the domain are underlined; variable residues among mammalian prion proteins⁹ are marked with asterisks. Line (1), mutations in human PrP that have been associated with inherited prion diseases⁷ (a stop codon at residue 145, which has been reported in addition to these point mutations⁷, is not considered here, nor is the Met232Arg mutation⁷, which is not contained in mature PrP^C). All of these residues are identical in wild-type human and mouse PrP. The polymorphism at codon 129 in human PrP, where homozygosity appears to increase susceptibility to sporadic CJD⁷, is marked by italics. Line (2), residues in PrP(121–231) for which experimental evidence has been presented that they contribute to the species barrier of prion disease transmission between mice and humans²⁰. **b**, Locations of selected residues in the three-dimensional structure of PrP(121–231). The backbone is shown in grey and the orientation of the molecule is as in Fig. 1c. The side chains of amino-acid residues with mutations that have been associated with inherited prion diseases⁷ are highlighted in red (line (1) in **a**). The solvent-accessible glycosylation sites at Asn 181 and Asn 197 are shown in green, and the disulphide bond is shown in yellow. Five residues that may be involved in the species barrier (line (2) in **a**) are shown in blue. Figure generated using the program MOLMOL²⁴.

proteins with folds similar to PrP(121–231), and the relative orientation of the three helices in PrP(121–231) is clearly different from the proposed four-helix-bundle model⁸.

Mapping onto the three-dimensional structure of PrP(121–231) of sequence variability in mammalian prion proteins⁹, of residues important for the species barrier of prion disease transmission²⁰ and for predisposition to familial prion diseases⁷, and of biochemical properties of the prion protein^{6,10} (Fig. 3a, b) shows the following. (1) Invariant residues in mammalian prion protein sequences are not clustered within the regions of regular secondary structure, but form an important part of the hydrophobic core. (2) The two glycosylation sites at Asn 181 and Asn 197 and the solvent-exposed, single Trp 145 are located on the negatively charged surface of the protein. (3) All six residues of PrP(121–231) for which mutation is believed to be associated with inherited prion diseases or predisposition to prion diseases are located in regular secondary structure elements or immediately adjacent to them, but none of these residues is located in the relatively isolated first helix. Three of these residues are part of the hydrophobic core, and three are located on the surface. They may therefore either destabilize the three-dimensional protein structure, or influence its ligand-binding properties. (4) As is generally observed for functionally related proteins from different

species, residues in PrP(121–231) that are variant in mammalian prion protein sequences⁹ are solvent accessible. (5) The disulphide bond 179–214 is highly shielded from solvent contact in the core of the protein.

It has been proposed that the species barrier of prion disease transmission between mice and humans is caused by an altered PrP^{Sc} binding site in PrP^C, which involves residues from the segment 96–167 (ref. 20). There are eight sequence differences between mouse and human PrP within this segment, of which five are contained in PrP(121–231). Four of these differences are located within or adjacent to the first helix, which might thus be part of a single binding site for PrP^{Sc} (Fig. 3a, b). The dipolar character of PrP(121–231) (Fig. 1b, c) might stabilize an orientation of PrP^C with its positively charged surface, which also includes hydrophobic surface patches, towards the cell membrane. Both glycosylation sites, as well as the aforementioned species barrier-related potential binding site for PrP^{Sc}, would then be located on the opposite, negatively charged surface. Further to these initial observations on possible structure–function correlations, we believe that the NMR structure of PrP(121–231) will provide a basis for more rational design of future *in vitro* and *in vivo* experiments on prion proteins and prion diseases. □

Received 13 May; accepted 5 June 1996.

- Griffith, J. S. *Nature* **215**, 1043–1044 (1967).
- Prusiner, S. B. *Science* **252**, 1515–1522 (1991).
- Weissmann, C. *Trends Cell Biol.* **4**, 10–14 (1994).
- Weissmann, C. *Nature* **375**, 628–629 (1995).
- Pan, K.-M. et al. *Proc. natn. Acad. Sci. U.S.A.* **90**, 10962–10966 (1993).
- Hornemann, S. & Glockshuber, R. *J. molec. Biol.* (in the press).
- Prusiner, S. B. *Arch. Neurol.* **50**, 1129–1153 (1993).
- Huang, Z. et al. *Proc. natn. Acad. Sci. U.S.A.* **91**, 7139–7143 (1994).
- Schätzl, H. M., Da Costa, M., Taylor, L., Cohen, F. E. & Prusiner, S. B. *J. molec. Biol.* **245**, 362–374 (1995).
- Stahl, N. & Prusiner, S. B. *FASEB J.* **5**, 2799–2807 (1991).
- Collinge, J. et al. *Nature* **370**, 295–297 (1994).
- Lledo, P.-M. et al. *Proc. natn. Acad. Sci. U.S.A.* **93**, 2403–2407 (1996).
- Sakaguchi, S. et al. *Nature* **380**, 528–531 (1996).
- Tobler, I. et al. *Nature* **380**, 639–642 (1996).
- Fischer, M. et al. *EMBO J.* **15**, 1255–1264 (1996).
- Wüthrich, K. *NMR of Proteins and Nucleic Acids* (Wiley, New York, 1986).

- Huang, Z., Prusiner, S. B. & Cohen, F. E. *Fold. Design* **1**, 13–19 (1996).
- Bernstein et al. *J. molec. Biol.* **112**, 535–542 (1977).
- Holm, L. & Sander, C. *Proteins* **19**, 165–173 (1994).
- Telling, G. C. et al. *Cell* **83**, 79–90 (1995).
- Güntert, P., Braun, W. & Wüthrich, K. *J. molec. Biol.* **217**, 517–530 (1991).
- Cornell et al. *J. Am. chem. Soc.* **117**, 5179–5197 (1995).
- Honig, B. & Nicholls, A. *Science* **268**, 1144–1149 (1994).
- Koradi, R., Billeter, M. & Wüthrich, K. *J. molec. Graph.* **14**, 51–55 (1996).
- Studier, F. W. & Moffat, B. A. *J. molec. Biol.* **189**, 113–130 (1986).
- Cavanagh, J., Fairbrother, W. J., Palmer, A. G. & Skelton, N. J. *Protein NMR Spectroscopy, Principles and Practice* (Academic, San Diego, 1996).

ACKNOWLEDGEMENTS. We thank M. Fischer, A. Raeber, C. Mumenthaler, and C. Weissmann for discussions, and the Centro Svizzero di Calcolo Scientifico for use of the NEC SX-3 and Cray J-90 computers. This work was supported by the Schweizerischer Nationalfonds and the ETH Zürich. S.H. was supported by the Boehringer-Ingelheim-Fonds.

CORRESPONDENCE and requests for materials should be addressed to R.G. (e-mail: rudi@mol.biol.ethz.ch).

Structural basis of RNA folding and recognition in an AMP–RNA aptamer complex

Feng Jiang*, R. Ajay Kumar*, Roger A. Jones† & Dinshaw J. Patel*

* Cellular Biochemistry and Biophysics Program, Memorial Sloan-Kettering Cancer Center, New York, New York 10021, USA

† Department of Chemistry, Rutgers, The State University of New Jersey, Piscataway, New Jersey 08855, USA

THE catalytic properties of RNA^{1,2} and its well known role in gene expression and regulation are the consequence of its unique solution structures. Identification of the structural determinants of ligand recognition by RNA molecules is of fundamental importance for understanding the biological functions of RNA, as well as for the rational design of RNA sequences with specific catalytic activities^{3–6}. Towards this latter end, Szostak *et al.*^{7,8} used *in vitro* selection techniques to isolate RNA sequences ('aptamers') containing a high-affinity binding site for ATP, the universal currency of cellular energy, and then used this motif to engineer ribozymes with polynucleotide kinase activity. Here we present the solution structure, as determined by multidimensional NMR spectroscopy and molecular dynamics calculations, of both uniformly and specifically ¹³C-, ¹⁵N-labelled 40-mer RNA containing the ATP-binding motif complexed with AMP. The aptamer adopts an L-shaped structure with two nearly orthogonal stems, each capped proximally by a G–G mismatch pair, binding the AMP ligand at their junction in a GNRA-like motif.

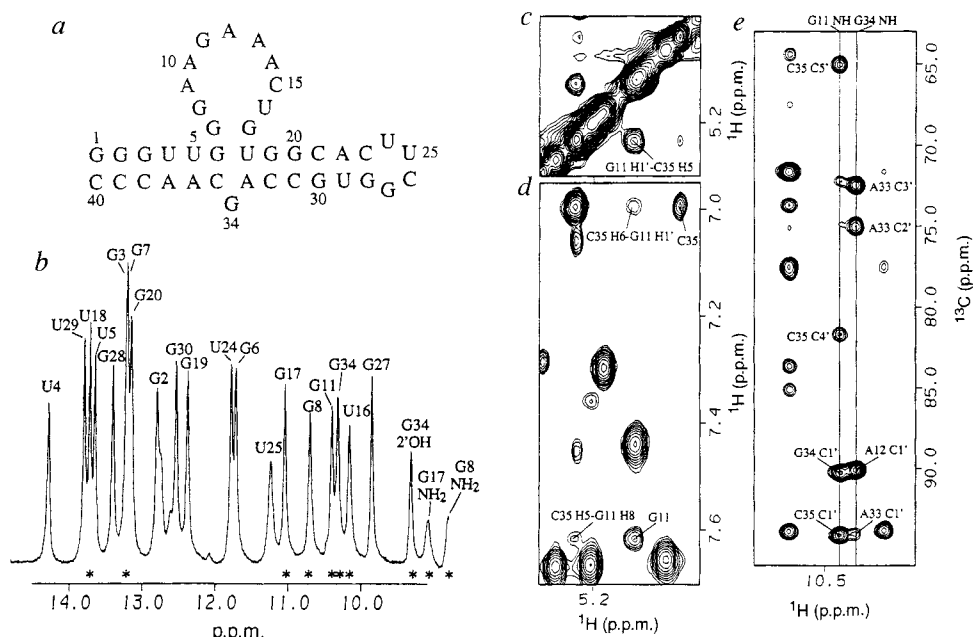
A 40-nucleotide RNA containing a consensus ATP-binding motif⁷ (Fig. 1a) was complexed with AMP, an analogue of ATP that binds to the RNA with a similar affinity. The AMP–RNA

aptamer complex gives well-resolved NMR spectra (Fig. 1b) that are amenable to structural characterization. Solution structures of the AMP–RNA complex (Fig. 2d) were calculated using a distance geometry and restrained molecular dynamics approach, guided by data from multidimensional NMR experiments (Table 1). The two stems of the RNA are orientated almost orthogonally ($106 \pm 9^\circ$) to each other, with the 11-residue purine-rich ATP-binding loop and the opposing bulged guanine at position 34 (G34) connecting the two (Fig. 2a). The ATP-binding loop (G7–G17) folds into two distinct hairpins which, together with the bulged G34, form a network of base pairing and stacking interactions that enclose the AMP within the binding pocket (Fig. 2b).

The first two purines of the loop, G7 and G8, continue to stack over the 5' strand of the left stem, leading to a hairpin turn spanning the G8–A10 residues (Fig. 2c, 3a). Base stacking is interrupted between residues G8 and A9, enabling the hairpin turn to form. Residues A9 to G11 are stacked in a right-handed helical fashion, with the purine base of AMP intercalated between A10 and G11. G11 closes the hairpin by forming a reverse Hoogsteen mismatched pair with G7 at the start of the ATP-binding loop (Fig. 3c) and, as evidenced by the nuclear Overhauser effect (NOE) data (Fig. 1c–e), extends the base stacking through C35 into the 3' strand of the left stem (Fig. 3b).

The second hairpin turn of the ATP-binding site forms over the right stem, where residues A17 and G34 form a Hoogsteen mismatched pair which is facilitated by a *syn* conformation of the glycosidic bond (χ) of G34 (Fig. 3c). Base stacking of the 5' strand of the right stem continues further into the second hairpin in a right-handed helical fashion from G17 to A13, leaving a discontinuity in the stacking at the A12–A13 step. Residues A13–C15 do not interact with AMP and constitute the least structured region of the ATP-binding motif. A12 stacks over G34, as an extension of the 3' strand of the right stem, and is held in place by hydrogen bonds to G34 and AMP (Fig. 3b). The 2' hydroxyl proton of G34, which hydrogen bonds to the N7 of A12, has a distinct downfield chemical shift (9.34 p.p.m.) and is readily detectable in

FIG. 1 a, The sequence and the predicted secondary structure of the 40-nucleotide ATP-binding RNA aptamer. b, The exchangeable-proton NMR spectrum of the AMP–RNA aptamer complex (20% excess AMP) in H₂O at 0 °C. All imino protons (except that of G1), one hydroxyl proton, and two amino protons are observed between 8.8 and 14.8 p.p.m. Assignments are shown over the resonances in the spectrum. Asterisks designate the resonances that appear on addition of AMP. c and d, Expanded regions of the two-dimensional (2D) NOESY spectrum ($\tau = 120$ ms) of the complex in ²H₂O at 15 °C, showing the NOEs between non-exchangeable protons of G11 and C35. Intraresidue H1'–H8/H6 peaks of G11 and C35 are labelled by residue. e, Expanded region of the 2D (¹H, ¹³C) HMQC–NOESY spectrum ($\tau = 150$ ms) of AMP(¹³C, ¹⁵N)-labelled RNA aptamer complex (50% excess AMP) in H₂O at 0 °C, showing NOEs between G11 imino proton and C35 sugar protons, and between G34 imino proton and A12 and A33 sugar protons. The ATP-binding RNA aptamers were synthesized by *in vitro* transcription from a DNA template with commercially available NTPs (nucleoside triphosphates) or (¹³C, ¹⁵N)-labelled NTPs using T7 RNA polymerase²⁴ and purified by gel electrophoresis. The concentration of AMP–RNA aptamer samples for NMR experiments was ~2 mM for unlabelled samples and ~3 mM for the (¹³C, ¹⁵N)-labelled sample in 10 mM sodium phosphate and 0.2 mM EDTA at pH 6.7. NMR spectra were acquired either on Varian Unityplus



600 or 500 spectrometers. Complete resonance assignment of the system was accomplished from analysis of 2D and 3D NMR spectra²⁵ using an unlabelled RNA–AMP sample, a (¹³C, ¹⁵N)-labelled RNA complexed with unlabelled AMP, or an unlabelled RNA complexed with (¹³C, ¹⁵N)-labelled AMP. All 1D and 2D spectra were processed using Varian VNMR software and the 3D spectra were processed using NMRpipe²⁶.

the spectrum of the complex (Fig. 1b).

The intercalation of the purine ring of the AMP between A10 and G11 of the first hairpin turn is of particular significance in that there is a striking similarity between the structure of this region of the complex and the GNRA motif⁹, a well characterized^{10–14}, extremely stable RNA secondary-structure element. Residues G8–A10 function as the first three residues of the GNRA-like

motif, with the characteristic unstacking and chain reversal between G8 and A9, whereas the AMP ligand comprises the fourth residue (Fig. 3a). The non-covalently attached AMP is oriented antiparallel to its stacked neighbours, A10 and G11, and parallel to its base-pairing partner, G8. This orientation of the ligand aligns its Watson–Crick edge towards the minor-groove edge of G8, leading to the hydrogen bonds shown in Fig. 3d, in

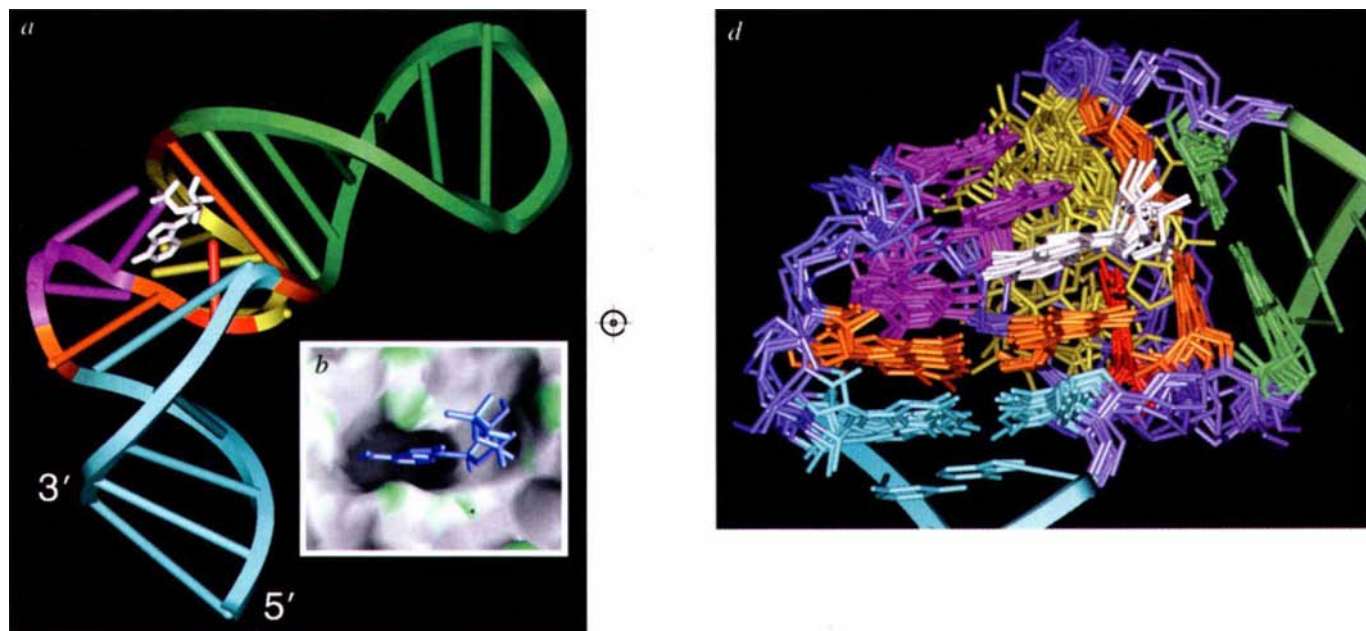
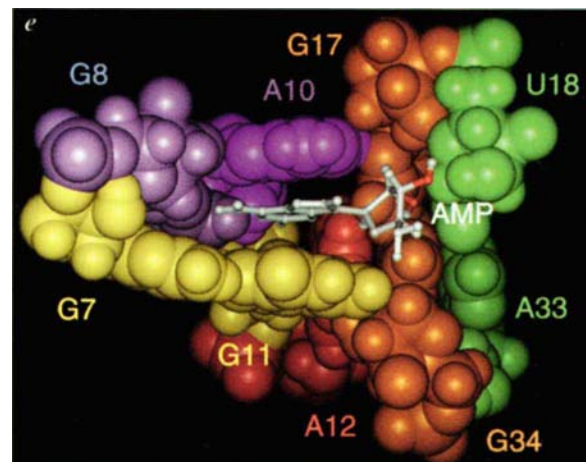


FIG. 2 *a*, Ribbon representation of the solution structure of the 40-mer RNA aptamer complexed with AMP. The bound AMP is shown in stick representation in white. The RNA ribbon is coloured to distinguish key residues in this and other colour figures. The left stem (G1–G6, C35–C40) is blue; the right stem–loop (U18–A33) is green. Residue pairs G7–G11, and G17–G34 form unusual mismatched pairs and are coloured orange. Residues G8–A10 constitute part of the GNRA-like structural motif and are coloured magenta. Residue A12 is red. The remainder of the large 11-residue RNA loop is less well structured and these residues, A13–U16, are shown in yellow. The AMP molecule is white. The angle between the two helical stems ($106^\circ \pm 9^\circ$) and the fold of the ATP-binding loop into two distinct hairpins are evident. The RNA structures have been analysed using CURVES²⁷. *b*, Binding of AMP in the RNA pocket, rendered using GRASP²⁸. The accessible surface of the region of RNA around AMP is shown here, shaded for the curvature of the surface from green (convex) to grey (concave). The accessible surface was calculated with a probe radius of 1.4 Å. AMP is coloured blue and shown in stick representation. *c*, Stereo view of the ATP-binding motif complexed with AMP. The RNA backbone is shown in purple. *d*, Superposed view of 8 final-distance refined structures of the AMP–RNA aptamer complex. The superposition was performed on all heavy atoms of the well-structured core of the ATP-binding motif, namely G6–A12, U16–U18, A33–C35 and AMP, which are shown in stick representation. Residues A13–U16 (yellow) are noticeably less well structured than the rest of the ATP-binding motif. A-form stems of the structure that has lowest r.m.s.d. to the average of 8 structures are shown in ribbon representation. *e*, Interaction of AMP with RNA residues in the binding pocket. The 2' and 3' hydroxyls of AMP are coloured red and are seen pointing into the right stem. 5'-terminal phosphate groups are omitted for clarity in *d* and *e*.



contrast to the sheared G·A pair in a typical GNRA motif¹⁰. In addition, the second hairpin facilitates the binding of AMP by orienting A12 orthogonal to the bases of AMP and G11, such that its exocyclic amino group forms a hydrogen bond to N3 of AMP (distance 1.91 ± 0.11 Å, angle $134 \pm 16^\circ$) in the first hairpin (Fig. 3b). Furthermore, the minor groove of the right stem abuts the ribose of AMP (Fig. 2e), which is immobilized by hydrogen bonds to the 2' hydroxyl (to G34 N7, distance 2.43 ± 0.03 Å, angle $158 \pm 6^\circ$, or G17 N3, distance 2.12 ± 0.20 Å, angle $166 \pm 1^\circ$) and 3' hydroxyl (to U18 2' hydroxyl, distance 2.19 ± 0.54 Å, angle $151 \pm 35^\circ$) groups.

The arrangement of the purine rings in the stacked core of the complex establishes the nearly orthogonal orientation of the two stems. Figure 3b illustrates the purine-stacked core at the junction of the two stems. In particular, we observe stacking between non-adjacent residues G11 and C35, A12 and G34, and the accompanying unstacking at steps G11–A12, A12–A13 and G34–C35. Concurrently, ribose sugars of residues G11, A12 and G34 are puckered in the C2'-endo conformation, in contrast to usual C3'-endo puckers in A-form RNA, allowing the backbone distortion that accompanies this unstacking. This pattern of stacking alignments, together with the GNRA-like motif, offers a plausible explanation for the invariant concentration of purines in the active

site of the ATP-binding aptamers⁷.

In summary, the structure reveals several means by which the RNA aptamer recognizes and binds the AMP molecule, including base-mismatch pairing, stacking interactions, hydrogen bonding and, remarkably, integrating the ligand in an extremely stable GNRA-like hairpin fold in which the intercalated adenine base is involved in an A·G mismatch pair. The ribose moiety of the AMP fits snugly in the minor groove of the right stem, such that the γ -phosphate of ATP would be exposed and available for catalysis. Furthermore, knowledge of these interactions provides a structural basis for understanding the published chemical modification and footprinting experiments⁷. As nearly all of the atoms of the purine base and the 2' and 3' hydroxyls of the ribose sugar of the ligand are either hydrogen bonded (Fig. 3d) or in van der Waals contact with the RNA (Fig. 2e), methylation or other modifications at these positions would clearly impair binding. However, despite this plethora of interactions, in which the aptamer contacts nearly half of the total accessible surface area of the ligand (259 ± 22 Å² out of 507 ± 17 Å²; Fig. 2b), the C8 position of the purine ring (the site of attachment of the ligand to the solid support during affinity chromatography) and the 5'-phosphate moiety remain exposed to the solvent. The exposed phosphate group is well suited for the rational design of ribozymes in which

TABLE 1 Statistics of NMR data and structures*

Distance restraints		
RNA (G6–U18, A33–C35)		441
Intraresidue		205
Sequential ($ i - j = 1$)		130
Medium range ($2 \leq i - j \leq 4$)		36
Long range ($ i - j \geq 5$)		70
AMP		7
RNA–AMP (G6–U18, A33–C35, AMP)		45
Structure statistics		
NOE violations		
Number > 0.2 Å		14.2 ± 2.4
Largest violation (Å)		0.48 ± 0.12
R.m.s.d. of violations (Å)		0.091 ± 0.015
Deviations from ideal covalent geometry		
Bond length (Å)		0.008 ± 0.002
Bond angle ($^\circ$)		3.1 ± 0.5
Improper ($^\circ$)		0.78 ± 0.24
Pairwise r.m.s.d. amongst 8 final structures (Å)	Base and sugar†	All heavy‡
ATP-binding motif + AMP		
G6–U18, A33–C35, AMP	2.00 ± 0.60	2.37 ± 0.62
Well-structured core + AMP		
G6–A12, U16–U18, A33–C35, AMP	1.48 ± 0.48	1.95 ± 0.55
GNRA-like motif		
G7–G11, AMP	1.49 ± 0.55	2.04 ± 0.64

Interproton distances were obtained from crosspeak volumes in NOESY spectra of the complex in H₂O (mixing time $\tau = 100$ ms) and ²H₂O ($\tau = 120$ ms) solution. Distance restraints were imposed with bounds at $\pm 20\%$ of the calculated distance, which were calibrated using pyrimidine H5–H6 as a reference. Non-stereospecific assignments were treated with r^{-6} averaging. In addition, crosspeaks that were not amenable to such quantitative analysis were manually classified into three categories, with corresponding distances and bounds: strong ($2.5, -0.7/+0.8$ Å), medium ($3.5, -1.5/+1.0$ Å), weak ($4.5, -2.5/+1.5$ Å). Additional distance restraints were obtained from quantitative analysis of crosspeak heights in the three-dimensional (3D) NOESY–HMQC ($\tau = 120$ ms) spectrum in ²H₂O. Distance restraints were imposed on Watson–Crick hydrogen-bond participants in the base pairs of the RNA stems. 1,000 initial structures of the complex were generated by a metric matrix distance geometry protocol using X-PLOR²³, which were quantitatively scored to identify 48 structures with least NOE violations, acceptable covalent geometry, and favourable van der Waals energy, in the ATP-binding motif. NMR data indicate that the two stems have a regular A-form conformation, hence on each of these distance-geometry embedded structures, the left (G1–G6, C35–C40) stem and right (U18–C23, G28–A33) stem structures were substituted with standard A-form duplexes and restrained with non-experimental distance, torsion and planarity restraints. The structures were subjected to three cycles of restrained molecular dynamics with a simulated annealing protocol, using the CHARMM force field with reduced charges²³. In each cycle of simulated annealing, the dynamics was initiated at 5 K, with a force constant of $2 \text{ kcal mol}^{-1} \text{ Å}^{-2}$ on all experimentally obtained distance restraints. The structure was heated to 1,000 K in 5 ps and allowed to evolve for 10 ps. Then the force constant on the experimentally obtained distance restraints was scaled up to $30 \text{ kcal mol}^{-1} \text{ Å}^{-2}$ in 8 ps. The system was allowed to evolve for another 10 ps, and then cooled slowly to 300 K in 14 ps and equilibrated for 20 ps. Snapshots of the structure taken every 0.5 ps for the last 4 ps were averaged and the resulting coordinates subjected to conjugate gradient minimization to a gradient of $0.1 \text{ kcal mol}^{-1} \text{ Å}^{-1}$. Eight final structures were chosen, based on the criterion of low restraint violation. Coordinates of the complex have been deposited with the Protein Data Bank (1ARA).

* All statistics and analysis unless otherwise indicated are restricted to the ATP-binding motif (G6–U18, A33–C35, AMP).

† Base and sugar heavy atoms.

‡ All heavy atoms.

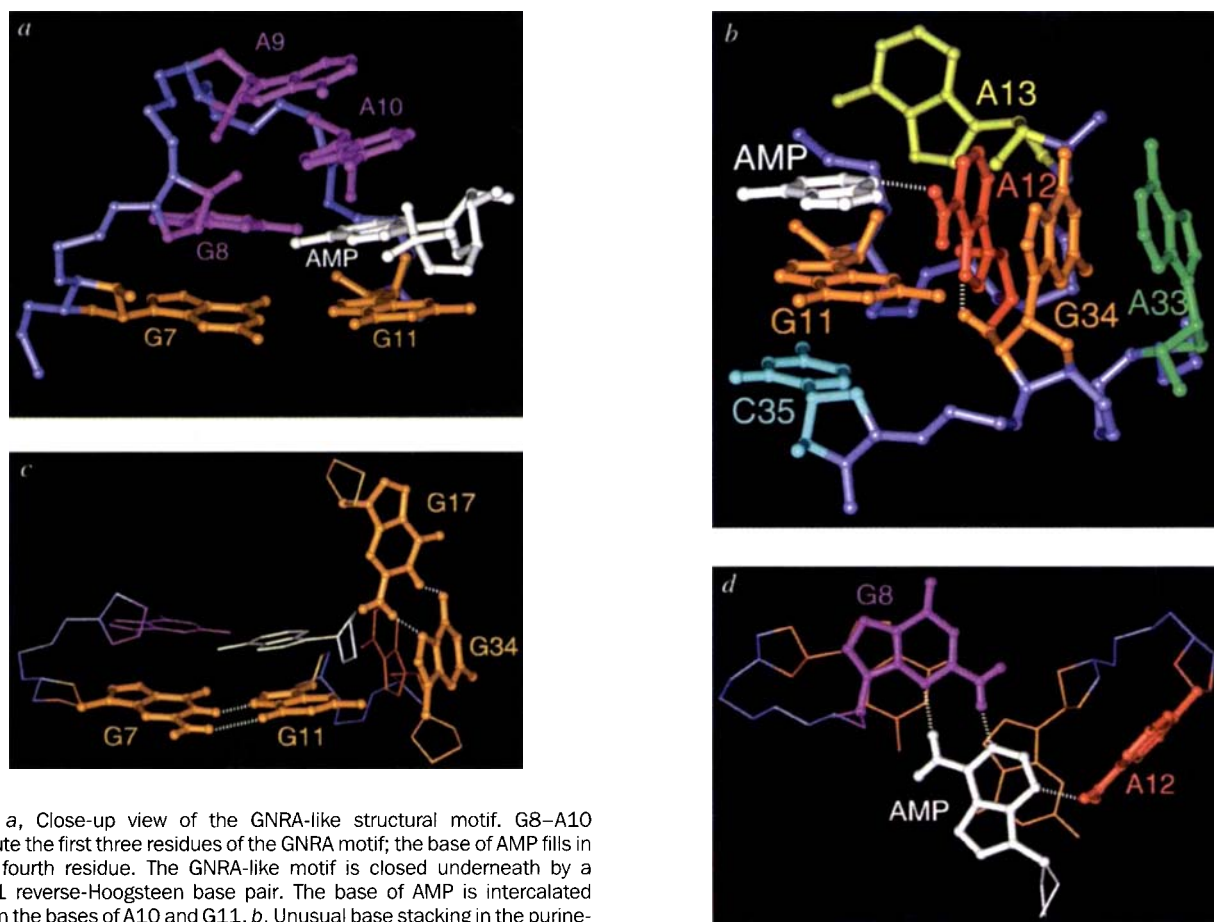


FIG. 3 *a*, Close-up view of the GNRA-like structural motif. G8–A10 constitute the first three residues of the GNRA motif; the base of AMP fills in as the fourth residue. The GNRA-like motif is closed underneath by a G7·G11 reverse-Hoogsteen base pair. The base of AMP is intercalated between the bases of A10 and G11. *b*, Unusual base stacking in the purine-stacked core of the ATP-binding motif. Hydrogen bonds from A12 to AMP and G34 are indicated with dashed lines. Some atoms are omitted for clarity. *c*, G7·G11 and G17·G34 mismatched Hoogsteen pairs at the junction of the stems and ATP-binding loop. Note that the glycosidic torsion angle (χ) of residue G34 is in the *syn* conformation. Hydrogen bonds are indicated by dashed lines. *d*, Unusual G8·AMP base pairing in the GNRA-

like structural motif. The minor groove edge of G8 pairs with the Watson–Crick edge of the AMP, using G8 N3 to AMP NH₂, G8 NH₂ to AMP N1 hydrogen bonds, and A12 NH₂ forms a hydrogen bond with AMP N3. Colour figures (except Fig. 2*b*) were prepared using Insight (Biosym/MSI).

the availability of a high-energy phosphodiester bond for driving catalysis is paramount. The recognition of AMP by RNA involves adaptive binding^{15–21}, whereby the AMP induces a conformational transition in the ATP-binding site on the RNA. In the absence of the ligand, the loop and bulged purines are exposed to the solvent, as evidenced by the absence of loop-guanine imino resonances in the spectrum of free RNA. The assembly of two hairpins and a

bulged guanine that interact to enclose the AMP within is likely to be initiated by the formation of a GNRA-like motif on addition of AMP. This solution structure of the AMP–RNA aptamer complex suggests that RNA molecules may facilitate catalysis by an ‘induced fit’ mechanism²² similar to that adopted by protein enzymes. □

Received 18 March; accepted 4 June 1996.

- Cech, T. R., Zaug, A. J. & Grabowski, P. J. *Cell* **27**, 487–496 (1981).
- Kole, R. & Altman, S. *Biochemistry* **20**, 1902–1906 (1981).
- Joyce, G. F. *Gene* **82**, 83–87 (1989).
- Robertson, D. L. & Joyce, G. F. *Nature* **344**, 467–468 (1990).
- Ellington, A. & Szostak, J. W. *Nature* **348**, 812–822 (1990).
- Tuerk, C. & Gold, L. *Science* **249**, 505–510 (1990).
- Sassanfar, M. & Szostak, J. W. *Nature* **364**, 550–553 (1993).
- Lorsch, J. R. & Szostak, J. W. *Nature* **371**, 31–36 (1994).
- Woese, C. R., Winker, S. & Gutell, R. R. *Proc. natn. Acad. Sci. U.S.A.* **87**, 8467–8471 (1990).
- Heus, H. A. & Pardi, A. *Science* **253**, 191–194 (1991).
- Varani, G., Cheong, C. & Tinoco, I. Jr *Biochemistry* **30**, 3280–3289 (1991).
- Orita, M. et al. *Nucleic Acids Res.* **21**, 5670–5678 (1993).
- Pley, H. W., Flaherty, K. M. & McKay, D. B. *Nature* **372**, 68–74 (1994).
- Scott, W. G., Finch, J. T. & Klug, A. *Cell* **81**, 991–1002 (1995).
- Puglisi, J. D. et al. *Science* **257**, 76–80 (1992).
- Aboul-ela, F., Kam, J. & Varani, G. *J. molec. Biol.* **253**, 313–332 (1995).
- Battiste, J. L., Tan, R., Frankel, A. D. & Williamson, J. R. *J. Biomol. NMR* **6**, 375–389 (1995).
- Puglisi, J. D., Chen, L., Blanchard, S. & Frankel, A. D. *Science* **270**, 1200–1203 (1995).

- Ye, X., Kumar, R. A. & Patel, D. J. *Chem. Biol.* **2**, 827–840 (1995).
- Fan, P. et al. *J. molec. Biol.* **258**, 480–500 (1996).
- Allain, F. H. T. et al. *Nature* **380**, 646–650 (1996).
- Koshland, D. E. Jr *Harvey Lect.* **65**, 33–57 (1971).
- Brünger, A. T. *X-PLOR, A System for X-ray Crystallography and NMR* (Yale University Press, New Haven, Connecticut, 1992).
- Nikonowicz, E. P. et al. *Nucleic Acids Res.* **20**, 4508–4513 (1992).
- Pardi, A. *Meth. Enzym.* **261**, 350–380 (Academic, New York, 1995).
- Delaglio, F. et al. *J. Biomol. NMR* **6**, 277–293 (1995).
- Lavery, R. & Sklenar, H. *J. biomolec. Struct. Dyn.* **6**, 63–91 (1988).
- Nicholls, A., Sharp, K. A. & Honig, B. *Proteins Struct. Funct. Genet.* **11**, 281–296 (1991).

ACKNOWLEDGEMENTS. F.J. and R.A.K. contributed equally to this work. We thank J. W. Sam for helpful comments on the manuscript. F.J. received a postdoctoral fellowship from the NIH. R.A.K. is a postdoctoral fellow of the Miriam and Benedict Wolf Cancer Research Fund. This work was supported by start-up funds to D.J.P.

CORRESPONDENCE and requests for materials should be addressed to D.J.P. (e-mail: pateld@mskcc.org).

Scientific showroom

Time to think ahead to the new academic year and items that may be needed, such as a system for cultivating hybridoma cells, a mass detector for liquid chromatographs and an accounting program to keep track of expenses.

JOUAN has introduced the MR22i, its first **multifunctional centrifuge** that can change from 4 × 50-ml conical cell culture tubes (as low as 500 r.p.m. in a swing-out rotor) to 8 × 50 ml at 21,255g (angle rotor) in just a few seconds. The rotor mounting system is responsible for this versatility. For high-speed centrifugation, the MR22i offers a choice of rotors, including drum and angle, as well as several sealed models, for tube volumes ranging from 0.25 ml to 100 ml. For high-performance centrifugation the company offers all volumes (18,516g to 36,668g). In low-speed centrifugation, 15- and 50-ml culture tubes, standard tubes (up to 180 ml) and even microtitre plates can be accommodated. To make the most of all these facilities, the new VideoSet operator dialogue system has been developed. The graphical screen on the MR22i is designed to make the setting of all parameters simple, which should help run-to-run reproducibility, as well as giving access to the database of rotor parameters.

Reader Enquiry No. 100

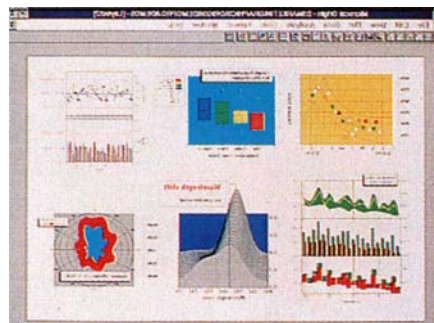
Software and hardware

Workstation Source has launched the FW5000, a rugged **fieldwork notebook computer** with built-in expansion and upgradability. It can withstand shocks of 100g, and is ideal for 'on-the-move' applications, such as field work and service, engineering and emergency services, as well as everyday situations where high processing power interfacing capabilities and durability are necessary. A number of processor options are offered, including 486DX4/100 and 133 MHz Pentium and beyond. Upgradability is said to be ensured through the new 'technology module', which allows for a quick and easy upgrade to the latest processors on the market, such as the Pentium Plus. Expandability is also simple through the four built-in expansion bays — options include CD drives, external hard drives, PCMCIA devices, batteries, power supplies and floppy drives. Like its larger counterpart, the FW7000 portable, the new notebook uses aerospace technology to meet MIL 810C shock and vibration specifications. Its low weight (4.54 kg) is achieved by using magnesium and aluminium components in a high strength spaceframe/skin structure. The display is a 26.4-cm dual-scan colour LCD as standard

with options of a TFT or a sunlight-readable transreflective monochrome display. An 85-key QWERTY keyboard with full-size keys is provided, complete with a mousepad, which can also be used as a signature/drawing pad. Both the keyboard and mousepad are sealed to protect against weather, dust and moisture. A 540-Mb hard disk is standard and can be expanded to 2.6-gigabyte fixed or removable. An unusual — but useful — feature is the standard 16-bit stereo card and stereo speakers. Also provided as standard are two serial ports and one parallel port. An integral NiMH battery provides up to three hours of use at full power (42 W). There is room for a second battery, which will provide six hours of power in total.

Reader Enquiry No. 101

Microcal Software has announced the release of Origin version 4.1 **graphics and data analysis software**. Along with a large selection of easy-to-use technical graphics and data-analysis tools, the program offers new ternary diagrams and bubble charts, a new calibration template, improved sigmoidal fit and FFT smoothing, and extended development support



Origin graphics and data-analysis tools.

for LabTalk script users. The software has been released in both 16-bit and 32-bit formats, providing speed and utilities, such as long file names for Windows95 users, while still offering a native version for Windows 3.1 users. For all registered users who currently have Origin 4.0, the company will automatically send the 4.1



Fieldwork computer.

upgrade for free. The novice user should find the program easy to learn, being able to enter, graph and analyse data quickly, and present it in a multitude of graphing and plotting formats, which can be customized by the user. The advanced user can automate data analysis by creating scripts with the built-in programming language, LabTalk. The 16-bit local memory restriction has been removed and this version can now support a large amount of data and graphs within a project. A typical project may contain multiple worksheets, each with hundreds of columns of data and hundreds of graphs. The 32-bit version supports unrestricted data size and an unlimited number of graphs within a project. The new calibration template adds a new dimension in laboratory data analysis by allowing engineers and technicians to obtain directly calibration curves from the instrument data.

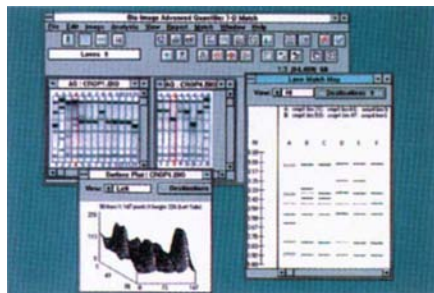
Reader Enquiry No. 102

Grant Manager 3.0 for Windows from Northern Lights Software is a **financial accounting program**, designed to act as a 'watchdog' over grant funds. This software features an array of functions to help organize and report the spending of research grant funds, contracts and departmental budgets. It helps track and retrieve information about expenditures, especially when working with several grants. The program should prove useful to financial managers who must monitor the accuracy of university accounting records. It can be used for both single and multiple grants, allowing users to enter orders, see up-to-the-minute balances, create custom reports, generate purchase orders and corroborate outside ledgers. There are said to be advantages over standard spreadsheets in purchase-order processing, searching and vendor archiving. It is designed with various pop-up menus, look-up lists, and automatic data entry validation. Each screen looks like an empty form. However, once the form is completed, specific function keys execute the various tasks. The program maintains running balances on all accounts, and provides a means to guard against overdraft and unspent funds left over from a particular budget period. This version includes a Windows interface, drop-down menus, switchable screen displays, on-line help, networking capability for multiple users, purchase order form designer and a custom report/query gener-

ator. This system is based on the Microsoft FoxPro xBase system and can be used in Novell, Lantastic, Novell Lite, Windows for Workgroups, Windows NT and other NetBios compatible networks.

Reader Enquiry No. 103

Bio Image has announced the release of a new family of one-dimensional **electrophoresis analysis software** for the PC and PowerMac, which includes the Advanced Quantifier (AQ) product line. The new line includes the AQ 1-D and the AQ 1-D Match. The system utilizes



Gel analysis and band-matching software.

proprietary algorithms for both programs to find lanes and bands automatically. The whole-band algorithm quantifies the band, providing results that are stated to be accurate and reproducible, and which can be compiled into reports customized

by the user. AQ 1-D is said to be able to match over a thousand lanes, displaying results in matrix or table form, as well as in dendrograms and contig maps. This matching capability can be used for forensic and paternity studies, RFLP, STRs, and other genetic mapping and typing applications. AQ accepts any TIFF file, and can drive most TWAIN-compatible scanners. The company also offers hardware systems for image acquisition that can be paired with AQ software to provide a complete system solution.

Reader Enquiry No. 104

Adept Scientific's DasyLab version 3 is a multifunctional Windows-based **data-acquisition program**, which is designed to be compatible with many popular data acquisition boards on the market. Designed for easy setup and use, it is said to allow complete control of each step, from data acquisition right through to the final printed report. The program simplifies data logging and display applications, as well as complex, high-speed acquisition and frequency domain analysis. It offers a wide choice of icons representing data inputs, outputs, controls, displays and functions for data reduction, signal analysis, and mathematical and statistical routines. These are connected on-screen to create a customized data-acquisition application, without programming. The new productivity enhancements include extended trigger functions to

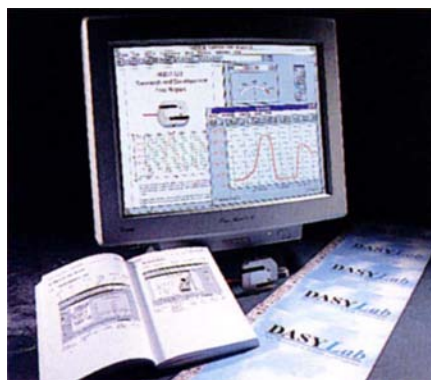
data clearly displayed on the printed page. Both RS232 and DDE support are upgraded in this new version, but for higher speed communication, a new option enables fast data acquisition from remote PCs using standard network components. DasyLab is available in both an extended edition (which includes VITool) and a basic version. Several optional extensions are also available, including tool kits for developing new functional modules. These allow the development of application-specific functions within the program's existing acquisition and display structure.

Reader Enquiry No. 105

Cultural awareness

Heraeus has recently introduced miniPerm, a new system designed for the simple and economical **cultivation of hybridoma cells** for the production of monoclonal antibodies. According to the manufacturer, upwards of 70 mg can be produced in 1-4 weeks. Downstream processing is said to be straightforward as the expressed antibody is contained in only 35 ml of medium. Accompanying this cultivation system is proSystem medium, which is optimized for use with the miniPerm system. This medium is said to perform better than serum-supplemented medium in terms of antibody production, requiring less than two per cent of the serum needed when using standard medium. For scientists wishing to work under completely protein-free conditions, proFree mab is offered as an optimized medium supplement for the cultivation of hybridoma and melanoma cells. Protein-free cultivation should provide a low background for the detection of cellular products and minimizes the risk of contamination with viruses and mycoplasma. The company states that pretesting of serum batches is not required and cultures are not subject to the characteristic biological fluctuations of serum-derived products. It is supplied sterile in 100-ml and 1,000-ml volumes.

Reader Enquiry No. 106



DasyLab data acquisition program.

make it easier to define pre- and post-trigger data, allowing users to specify the area of interest to be captured. A new equation editor allows maths expressions to be entered directly into the program's maths module for continuous, real-time computation. Its new VITool provides a means of visualizing and documenting a process or test rig. The VITool displays icons and can be combined with bitmap images, graphical objects, control buttons, switches and text to create screens that mimic the entire process, allowing the process design to be visualized on-screen. Users can then define the layout of a QC or test report, with the relevant

Shandon's Cryostat range has been extended with the addition of the Cryotome SME, a motorized electronic unit that is designed to provide **trouble-free sectioning of frozen specimens**. The new unit features electronic automated, specimen advance/retract and independent specimen temperature control down to -35°C , together with three-axis specimen orientation. It offers the user a choice of motorized or manual cutting with hands-free sectioning controlled through a touch key or foot switch. The Cryotome SME accommodates a choice of knife holders for use with disposable blades or conventional knives. By having the microtome situated outside the cryo

ADVERTISEMENTS



Gramsch Laboratories

- Fast Peptide Synthesis: weekend setup:
25 μmoles : 10 AA DM 650.00; 15 AA
DM 730.00; 20 AA DM 780.00
- Antibody Production Service:
Package : DM 2100.00

New Antisera:

G-Protein subunit : $\alpha 12$; $\alpha 13$; $\alpha 14$; $\beta 2$; $\gamma 5$
Nociceptin
Somatostatin Rec. mSSTR 1/2A/2B/3

Phone 08138 1769 or 08131 5679-0
Fax 08138 1260 or 08131 5679-24
WWW: ourworld.compuserve.com/
homepages/jgramsch/gi.htm

Gramsch Laboratories

D-85247 SCHWABHAUSEN, Germany

READER ENQUIRY NO. 11

OLIGONUCLEOTIDES

£1 per base

- stringent quality control
- express turnaround
- column purified: ready to use

By research labs for research labs

Molecular Biology Services

King's College School of Medicine and Dentistry
Rayne Institute, 123 Coldharbour Lane, London SE5 9NU
Tel: 0171-346 3126 Fax: 0171-733 3877

READER ENQUIRY NO. 10

chamber, the unit provides a large, clean working area, which is said to be easy to clean, fumigate and service.

Reader Enquiry No. 107

Bioseparations

Gynkosoftware is a PC-based software package from HPLC Technology for the **acquisition and control of chromatographic data**. It provides the essential information for compliance with GLP and other recognized quality standards. Designed to operate with any commonly used HPLC system, the software offers fast 32-bit processing and has a real-time multitasking capability, combined with a Windows interface. The package simultaneously acquires, processes and presents sample data from an unlimited number of HPLC systems, providing a fully traceable qualification of each analysis. Every event, manual or automatic, whether expected or unexpected, is recorded and stored in a real-time audit trail, complete with date, time stamp and operator identification. All information is documented and retained in a protected memory as an integral part of each sample analysis, which cannot be accessed or interfered with, ensuring the security necessary for regulatory compliance. Consistent with real-time data acquisition, Gynkosoftware can also be programmed to react intelligently to any fault conditions in individual instruments, triggering the appropriate hardware or software response, including alarm signals.

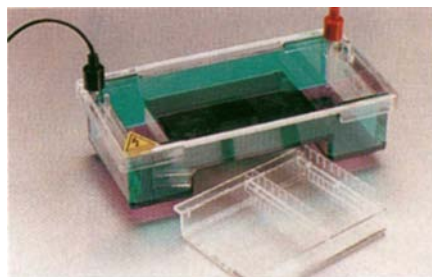
Reader Enquiry No. 108

Merck now offers BioBeads, **silica paramagnetic particles** for the adsorption and isolation of DNA and RNA from agarose, cell lysates or serum. The manufacturer states that using paramagnetic particles with a magnet is reliable and convenient and that without centrifugation, liquids can be easily removed from the bottom of the tube while the particles with the adsorbed nucleic acids are attracted by the magnet. A versatile magnetic rack is available for convenient magnetic separations. For frozen research budgets the company also offers a less expensive rectangular rare-earth magnet.

Reader Enquiry No. 109

Molecular Bio-Products has added colour to its line of **horizontal electrophoresis gel systems**. Both the Midi and Mega gel boxes, with 11-cm × 14-cm and 20-cm × 25-cm gel beds, respectively, are now available in a teal and purple colour scheme, as well as the

standard clear models. The new colours are intended to help differentiate and identify gel boxes for special procedures. All of the horizontal gel systems feature a darkened stage that clearly defines gels and facilitates sample loading. Transporting and analysing fragile gels is also said to be simplified with the easily



Molecular Bio-Products' new gel boxes.

removable ultraviolet-transparent gel tray. Combined with a gel-casting tray and comb system that is designed to ensure straight, even wells, these gel systems should make sample analysis quick and easy. The manufacturer states that these durable gel boxes are guaranteed for five years against leaking.

Reader Enquiry No. 110

Instrumental analysis

Kühnast now offers the MMSP 8000 **portable multispectrometer instrument** for production of high-resolution spectra. The small portable unit is designed to be easy to use, with the equipment linked to a notebook computer through a PCMCIA interface. The main optical component of this instrument is a product of Zeiss. It has an active wavelength range from 190 nm up to 780 nm, ultraviolet to visible, absorbance and emissions spectra. The optical injection connects through fibre waylines and an SMA plug. Now the instrument is prepared to include the software and other available Windows programs.

Reader Enquiry No. 111

Perkin-Elmer Sciex has introduced the API 100LC, a new **benchtop liquid chromatography mass detector** based on its atmospheric pressure quadrupole ionization (API) benchtop single quadrupole liquid chromatography/mass spectrometry (LC/MS) system. This mass detector is designed to provide detection capabilities for liquid chromatographers, providing access to mass information for unambiguous identification, together with mass and structural information right on the benchtop. The system provides complete control of an LC and autosampler from Perkin-Elmer and other manufacturers, and takes advantage of the capabilities of the company's new MExpress software, which is designed to enable novice users to analyse a sample with minimal involvement. All

system components (mass detector, LC, autosampler and data processing system) operate as a unified system with single-point control. MExpress software allows users to log a sample in, control the LC and autosampler, acquire the mass spectral information and produce pertinent reports — all automatically.

Reader Enquiry No. 112

Packard Instruments has announced the new TopCount HTS **microplate scintillation and luminescence counter** for high-throughput sample counting and analysis. This is the company's first microplate counter with the ability to count samples directly in 384-well microplates for radioisotopic and luminescence assays. The counter's newly designed 40-plate stacker cassette can load more than 15,000 samples at once for unattended counting. A new 12-detector luminescence counting option is said to enable the unit to count 10,000 or more luminescence samples per hour using reagent systems such as Packard's LucLite for measuring the luciferase reporter gene. This unit also counts samples in either the 384-well format or the standard 96-well format to ensure maximum sample processing versatility. Reagent costs can be reduced greatly with the 384-well microplate format, down to 50–80 microlitres per well. According to the manufacturer, robotic sample prepara-

ADVERTISEMENTS

NEWCASTLE UNIVERSITY
MOLECULAR BIOLOGY UNIT



PROTEIN SEQUENCING

Membrane-bound, solid or solution samples

£12 per cycle

5 cycles minimum - no set up charge

For further information contact Dr. Joe Gray

Phone: +44 191 222 8612

Fax: +44 191 222 8129

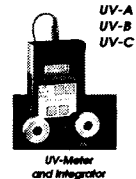
Email: JOE.GRAY@newcastle.ac.uk

READER ENQUIRY NO. 18

KÜHNAST
Strahlungstechnik
Poststraße 56
63607 Wöhrtenbach
phone: 0049 (0)6053-9650
fax: 0049 (0)6053-4636
GERMANY

UV-products worldwide

**Datalogger for UV-C
watersterilizer &
Ozon - control units**



UV-measuring systems

READER ENQUIRY NO. 52

tion and loading is also enhanced through the instrument's robotic interface.

Reader Enquiry No. 113

Hewlett-Packard has introduced the HP 5973 mass selective detector (MSD), a **benchtop mass spectrometer** that is designed to provide sensitivity, improved uptime, easy maintenance and productivity. The company has designed the instrument for a variety of analytical areas, including the analyses of pesticides, volatile and semivolatile organic compounds, toxic compounds in air and drug metabolites in bodily fluids. In the electron-ionization (EI) selected-ion-monitoring (SIM) mode, the instrument is said to provide femtogram-level sensitivity. For analyses performed in the EI full-scan mode, it can generate library-searchable spectra with one picogram of octafluoronaphthalene. In full-scan modes for positive chemical ionization (PCI) and negative chemical ionization (NCI), the system provides picogram-level sensitivity. Key sensitivity enhancing features of the HP 5973 MSD include the new analyser components; a high-energy dynode electron-multiplier detector that enhances gain and incorporates improved ion optics to reduce noise; a metalized quadrupole designed to improve resolution and sensitivity; a PCI/NCI option for classes of compounds, such as polyhalogenated samples, which respond well to NCI analyses; and a short capillary direct interface, together with the temperature-control mechanisms to minimize sample degradation and loss of chromatographic resolution. As a result, GC methods can be adapted more quickly to GC/MS. A new, independently heated ion source and quadrupole mass filter heat the ion source up to 250 °C, which helps keep the analyser cleaner for longer, even when dirty samples are analysed.

Reader Enquiry No. 114

The new BioLumin 960 **micro-assay reader** from Molecular Dynamics is designed to offer both fluorescence and absorbance microplate analysis capabilities and comes with a combination of features that make it suitable for both low- and high-throughput phases of micro-assay projects. The new reader features one-femtomole, fluorescent-dye sensitivity



Fluorescence and absorbance analysis.

and reads microplates in absorbance, fluorescence, or both modes simultaneously, in as little as twelve seconds. As the sample loading drawer closes, the microplates align for automatic reading. Moreover, the reading chamber is temperature-controlled from ambient to 45 °C. There is an optional RS232 interface, which enables synchronized robotic operation. Experiment software for Macintosh computers provides system control, data analysis and standardized result printouts. Typical applications for the BioLumin 960 micro-assay reader include apoptosis, cytotoxicity, cell viability, cytokine, ligand-receptor and toxicology assays.

Reader Enquiry No. 115

Labware

Mettler-Toledo has developed a system that allows **integration of balances and scales into laboratory systems**. The LocalCAN universal interface from Mettler-Toledo is built into all precision balances of the so-called professional and standard levels. The simple and economical basic-level balance can also be retrofitted with this system, if required. The system is not only compatible with the



Mettler-Toledo's LocalCAN interface.

standard RS232, but can also meet future interface standards. In contrast with a simple RS232 interface, the universal interface allows the simultaneous attachment of up to five peripheral devices to a balance. For example, a balance can simultaneously control a printer, receive the data of a bar-code reader for sample identification, as well as communicate bidirectionally with a PC or host of a LIMS. The compatibility among the different balance types is ensured by standardized interface commands. For the system planner, the modular concept should provide long-term consistency with the necessary flexibility.

Reader Enquiry No. 116

The new generation of the Finnpiptette BioControl from LabSystems is now a complete **liquid handling system** with one universal handle and ten interchangeable tip-cone modules. The ten-in-

one handle fits all five single-channel and five multichannel modules, allowing users to pipette at three speeds, without the need to purchase multiple units. A snap-lock mechanism makes it fast and easy to attach and remove each module. Volumes range from 0.5 µl to 10 ml. Moreover, this new liquid handling system uses trigger-action pipetting and has been ergonomically designed with features such as Soft-Touch tip ejection, autoclavable tip-cone modules, simplified programming and automatic power cut-off and restart. It can be calibrated within the laboratory, using an easy-to-follow Windows-based software program.

Reader Enquiry No. 117

Literature

Biosynth has introduced an expanded **catalogue** that includes more than 300 enzyme substrates, enzyme inducers and culture media additives. This catalogue includes applications, enzyme-substrate reaction mechanisms and an enzyme-substrate cross-reference section. Information is provided for each chemical, including the chemical structure, the chemical formula, molecular weight, CAS number, chemical synonyms and abbreviations, as well as details on pricing and availability. Among the products that the company offers are luciferin, luciferase, coelenterazine, X-glcUA (BCIG), X-gal, MUG, IPTG, BCIP and other new substrates.

Reader Enquiry No. 118

These notes are compiled by Brendan Horton from information provided by the manufacturers. For more details, fill in the reader service card bound inside the journal.

ADVERTISEMENT

TaqEXPRESS

Nuclease deleted Taq DNA Polymerase (KlenTaq)
(e.g. PNAS 91 pp2216-2220)

- Greatly superior performance in primer extension reactions
- Stable at 99°C
- Patented and licensed (US 5,436,149)
- Inexpensive: £0.09/unit (sold in 500 and 5000 units)

To order: info@genpak.co.uk
Fax: +44 1273 704499 Tel: +44 1273 704470

GENPAK Limited

Science Park Square, Folmer, Brighton BN1 9SB, UK

READER ENQUIRY NO. 53